

**International Journal of Advances in
Engineering & Technology (IJAET)**

Anniversary Issue



Volume-2, Issue-1

URL : <http://www.ijaet.org>

E-mail : editor@ijaet.org

Table of Content

S. No.	Article Title & Authors (Vol. 2, Issue. 1, Jan-2012)	Page No's
1.	STATE OF ART: HAND BIOMETRIC Sarâh BENZIANE and Abdelkader BENYETTOU	1-9
2.	USING DYNAMIC DUAL KEYS ENCRYPTION ALGORITHM AS PARTIAL ENCRYPTION FOR A REAL-TIME DIGITAL VIDEO Abdul Monem S. Rahma and Basima Z.Yacob	10-18
3.	DESIGN AND PROTOTYPING OF A MINIATURIZED SENSOR FOR NON-INVASIVE MONITORING OF OXYGEN SATURATION IN BLOOD Roberto Marani, Gennaro Gelao and Anna Gina Perri	19-26
4.	EFFECTS OF PGPR ON GROWTH AND NUTRIENTS UPTAKE OF TOMATO Shahram Sharafzadeh	27-31
5.	THE APPLICATION OF PSO TO HYBRID ACTIVE POWER FILTER DESIGN FOR 3 PHASE 4-WIRE SYSTEM WITH BALANCED & UNBALANCED LOADS B. Suresh Kumar, K. Ramesh Reddy & S. Archana	32-42
6.	A SURVEY OF COUPLING MEASUREMENT IN OBJECT ORIENTED SYSTEMS V. S. Bidve and Akhil Khare	43-50
7.	THE COMPUTER ASSISTED EDUCATION AND ITS EFFECTS ON THE ACADEMIC SUCCESS OF STUDENTS IN THE LIGHTING TECHNIQUE AND INDOOR INSTALLATION PROJECT COURSE İsmail Kayri, Muhsin Tunay Gençoğlu and Murat Kayri	51-61
8.	FRACTAL CHARACTERIZATION OF EVOLVING TRAJECTORIES OF DUFFING OSCILLATOR Salau, T. A.O. and Ajide, O.O.	62-72
9.	SANKEERNA: A LINEAR TIME, SYNTHESIS AND ROUTING AWARE, CONSTRUCTIVE VLSI PLACER TO ACHIEVE SYNERGISTIC DESIGN FLOW Santeppa Kambham1 and Siva Rama Krishna Prasad Kolli	73-89
10.	A NEW VARIANT OF SUBSET-SUM CRYPTOSYSTEM OVER RSA Sonal Sharma, Saroj Hiranwal, Prashant Sharma	90-97

11. A COMPACT DUAL BAND PLANAR RMSA FOR WLAN/WIMAX APPLICATIONS 98-104
C. R. Byraredddy, N. C. Easwar Reddy, C. S. Sridhar
12. VLSI ARCHITECTURE FOR LOW POWER VARIABLE LENGTH ENCODING AND DECODING FOR IMAGE PROCESSING APPLICATIONS 105-120
Vijaya Prakash. A.M & K.S. Gurumurthy
13. VERIFICATION ANALYSIS OF AHB-LITE PROTOCOL WITH COVERAGE 121-128
Richa Sinha, Akhilesh Kumar and Archana Kumari Sinha
14. IMPACT OF VOLTAGE REGULATORS IN UNBALANCED RADIAL DISTRIBUTION SYSTEMS USING PARTICLE SWARM OPTIMIZATION 129-138
Puthireddy Umapathi Reddy, Sirigiri Sivanagaraju
15. STUDY ON PERFORMANCE OF CHEMICALLY STABILIZED EXPANSIVE SOIL 139-148
P. VenkaraMuthyalu, K. Ramu and G.V.R. Prasada Raju
16. DESIGNING AN AUTOMATED SYSTEM FOR PLANT LEAF RECOGNITION 149-158
Jyotismita Chaki and Ranjan Parekh
17. FUZZY CONTROL OF SQUIRREL CAGE INDUCTION MACHINE WIND GENERATION SYSTEM 159-167
B. Ravichandra Rao and R. Amala Lolly
18. AN ADVANCED WIRELESS SENSOR NETWORK FOR LANDSLIDE DETECTION 168-178
Romen Kumar.M & Hemalatha
19. EVALUATION OF PHONETIC MATCHING APPROACHES FOR HINDI AND MARATHI: INFORMATION RETRIEVAL 179-189
Sandeep Chaware and Srikantha Rao
20. DESIGN OF ENERGY-EFFICIENT FULL ADDER USING HYBRID-CMOS LOGIC STYLE 190-202
Mohammad Shamim Imtiaz, Md Abdul Aziz Suzon, Mahmudur Rahman
21. EXAM ONLINE: E-ENABLING EXTENDED LEARNING, ANSWER AND ESSAY EXAMINATIONS 203-209
Abdulghader. A. Ahmed, Dalbir S., Ibrahim M.

22. NOISE MODELING OF SIGE HBT BASED ON THE CHARACTERIZATION OF EXTRACTED Y- AND Z- PARAMETERS FOR HF APPLICATIONS 210-219
Pradeep Kumar and R.K. Chauhan
23. DIELECTRIC PROPERTIES OF NORTH INDIAN OCEAN SEAWATER AT 5 GHZ 220-226
A.S. Joshi, S.S. Deshpande, M.L.Kurtadikar
24. AN EFFICIENT DECISION SUPPORT SYSTEM FOR DETECTION OF GLAUCOMA IN FUNDUS IMAGES USING ANFIS 227-240
S.Kavitha, K.Duraiswamy
25. STEP-HEIGHT MEASUREMENT OF SURFACE FUNCTIONALIZED MICROMACHINED MICROCANTILEVER USING SCANNING WHITE LIGHT INTERFEROMETRY 241-248
Anil Sudhakar Kurhekar and P. R. Apte
26. EXPERIMENTAL INVESTIGATION ON FOUR STROKE CERAMIC HEATER SURFACE IGNITION C.I. ENGINE USING DIFFERENT BLENDS OF ETHYL ALCOHOL 249-257
R.Rama Udaya Marthandan, N.Sivakumar, B. Durga Prasad
27. PERFORMANCE VERIFICATION OF DC-DC BUCK CONVERTER USING SLIDING MODE CONTROLLER FOR COMPARISON WITH THE EXISTING CONTROLLERS - A THEORETICAL APPROACH 258-268
Shelgaonkar (Bindu) Arti Kamalakar, N. R. Kulkarni
28. PERFORMANCE EVALUATION OF DS-CDMA SYSTEM USING MATLAB 269-281
Athar Ravish Khan
29. RECENT PHILOSOPHIES OF AGC OF A HYDRO-THERMAL SYSTEM IN DEREGULATED ENVIRONMENT 282-288
L. ShanmukhaRao1, N.Venkata Ramana
30. DYNAMIC ROUTING SCHEME IN ALL-OPTICAL NETWORK USING RESOURCE ADAPTIVE ROUTING SCHEME 289-298
S. Suryanarayana, K.Ravindra, K. Chennakesava Reddy
31. ENHANCED BANDWIDTH UTILIZATION IN WLAN FOR MULTIMEDIA DATA 299-308
Z. A. Jaffery, Moinuddin, Munish Kumar
32. ANALYSIS AND INTERPRETATION OF LAND RESOURCES USING REMOTE SENSING AND GIS: A CASE STUDY 309-314

S.S. Asadi, B.V.T.Vasantha Rao, M.V. Raju and M.Anji Reddy

33. IPV6 DEPLOYMENT STATUS, THE SITUATION IN AFRICA AND WAY OUT 315-322
Agbaraji E.C., Opara F.K., and Aririguzo M.I.
34. STUDY AND REALIZATION OF DEFECTED GROUND STRUCTURES IN THE PERSPECTIVE OF MICROSTRIP FILTERS AND OPTIMIZATION THROUGH ANN 323-330
Bhabani Sankar Nayak, Subhendu Sekhar Behera, Atul Shah
35. ANALYSIS OF DISCRETE & SPACE VECTOR PWM CONTROLLED HYBRID ACTIVE FILTERS FOR POWER QUALITY ENHANCEMENT 331-341
Jarupula Somlal, Venu Gopala Rao Mannam
36. COMPARISONS AND LIMITATIONS OF BIOHYDROGEN PRODUCTION PROCESSES: A REVIEW 342-356
Karthic Pandu and Shiny Joseph
37. MORPHOMETRIC AND HYDROLOGICAL ANALYSIS AND MAPPING FOR WATUT WATERSHED USING REMOTE SENSING AND GIS TECHNIQUES 357-368
Babita Pal, Sailesh Samanta and D. K. Pal
38. ADAPTIVE HYSTERESIS BAND CURRENT CONTROL FOR TRANSFORMERLESS SINGLE-PHASE PV INVERTERS 369-376
B. Nagaraju , K. Prakash
39. PARAMETRIC STUDY OF A NOVEL STACKED PATCH ANTENNA 377-384
V. Rajya Lakshmi, M. Sravani, G.S.N.Raju
40. BIOMETRICS STANDARDS AND FACE IMAGE FORMAT FOR DATA INTERCHANGE - A REVIEW 385-392
Nita M. Thakare and V. M. Thakare
41. COMPARATIVE ANALYSIS OF LOW-LATENCY ON DIFFERENT BANDWIDTH AND GEOGRAPHICAL LOCATIONS WHILE USING CLOUD BASED APPLICATIONS 393-400
N. Ajith Singh and M. Hemalatha
42. EVALUATION OF TEXTURAL FEATURE EXTRACTION FROM GRM FOR PROSTATE CANCER TRUS MEDICAL IMAGES 401-409
R.Manavalan and K.Thangavel
43. ANALYSIS AND MULTINOMIAL LOGISTIC REGRESSION MODELLING OF WORK STRESS IN MANUFACTURING 410-418

INDUSTRIES IN KERALA, INDIA

K. Satheesh Kumar and G. Madhu

44. AUDIO DENOISING USING WAVELET TRANSFORM 419-425
B. JaiShankar and K. Duraiswamy
45. HYBRID ACTIVE POWER FILTER USING FUZZY DIVIDING FREQUENCY CONTROL METHOD 426-432
SaiRam.I, Bindu.V and K.K. Vasishta Kumar
46. MINIMUM LINING COST OF TRAPEZOIDAL ROUND CORNERED SECTION OF CANAL 433-436
Syed Zafar Syed Muzaffar, S. L. Atmapoojya, D.K. Agarwal
47. VOLTAGE CONTROL AND DYNAMIC PERFORMANCE OF POWER TRANSMISSION SYSTEM USING STATCOM AND ITS COMPARISON WITH SVC 437-442
Amit Garg and Sanjai Kumar Agarwal
48. ASSOCIATION RULE MINING ALGORITHMS FOR HIGH DIMENSIONAL DATA – A REVIEW 443-454
K.Prasanna and M.Seetha
49. ACHIEVING EFFICIENT LOAD BALANCING IN PEER TO PEER NETWORK 455-462
Ritesh Dayama, Ranjeet Kagade, Kedar Ghogale
50. ANALYSIS AND SIMULATION OF SERIES FACTS DEVICES TO MINIMIZE TRANSMISSION LOSS AND GENERATION COST 463-473
M. Balasubba Reddy, Y. P. Obulesh and S. Sivanaga Raju
51. MODELING AND SIMULATION OF THE PATCH ANTENNA BY USING A BOND GRAPH APPROACH 474-484
Riadh Mehouchi, Hichem Taghouti, Sameh Khmailia and Abdelkader Mami
52. DESIGN AND VERIFICATION ANALYSIS OF AVALON INTERRUPT INTERFACE WITH COVERAGE REPORT 485-492
Mahesh Kumar Jha, Richa Sinha and Akhilesh Kumar
53. CONCATENATION OF BCH CODE WITH SPACE TIME CODE A LOW SNR APPROACH FOR COMMUNICATION OVER POWER LINES FOR SUBSTATION AUTOMATION 493-500
Rajeshwari Itagi, Vittal K. P., U. Sripati
54. UTILIZATION OF EXTRUSION AS AN ADVANCED 501-507

MANUFACTURING TECHNIQUE IN THE MANUFACTURE OF
ELECTRIC CONTACTS

Virajit A. Gundale ,Vidyadhar M. Dandge

55. CONTROL AND PERFORMANCE OF A CASCADED H-BRIDGE MLI AS STATCOM 508-519
M. Vishnu Prasad and K. Surya Suresh
56. IMAGE RETRIEVAL USING TEXTURE FEATURES EXTRACTED USING LBG, KPE, KFCG, KMCG, KEVR WITH ASSORTED COLOR SPACES 520-531
H.B.Kekre, Sudeep D. Thepade, Tanuja K. Sarode, Shrikant P. Sanas
57. DEVELOPMENT OF A SIMPLE & LOW-COST INSTRUMENTATION SYSTEM FOR REAL TIME VOLCANO MONITORING 532-542
Didik R. Santoso, Sukir Maryanto and A.Y. Ponco Wardoyo
58. INTENSIFIED ELGAMAL CRYPTOSYSTEM (IEC) 543-551
Prashant Sharma, Amit Kumar Gupta, Sonal Sharma
59. A ZIGZAG-DELTA PHASE-SHIFTING TRANSFORMER AND THREE-LEG VSC BASED DSTATCOM FOR POWER QUALITY IMPROVEMENT 552-563
R.Revathi and J.Ramprabu
60. MODELING AND SIMULATION OF NANOSENSOR ARRAYS FOR AUTOMATED DISEASE DETECTION AND DRUG DELIVERY UNIT 564-577
S.M. Ushaa, Vivek Eswaran
61. MISSING BOUNDARY DATA RECONSTRUCTION BY AN ALTERNATING ITERATIVE METHOD 578-586
Chakir Tajani and Jaafar Abouchabaka
62. A COMPREHENSIVE PERFORMANCE ANALYSIS OF ROUTING PROTOCOL FOR ADHOC NETWORK 587-593
Sachin Dahiya, Manoj Duhan, Vikram Singh
63. COMPARISON OF GA AND LQR TUNING OF STATIC VAR COMPENSATOR FOR DAMPING OSCILLATIONS 594-601
Nuraddeen Magaji, Mukhtar F. Hamza, Ado Dan-Isa
64. INVERTED SINE PULSE WIDTH MODULATED THREE-PHASE 602-610

CASCADED MULTILEVEL INVERTER

R.Seyezhai

65. ON THE SUNSPOT TIME SERIES PREDICTION USING JORDON ELMAN ARTIFICIAL NEURAL NETWORK (ANN) 611-621

Rohit R. Deshpande and Athar Ravish Khan

66. SESSION DATA PROTECTION USING TREE-BASED DEPENDENCY 622-631

G. Shruthi, Jayadev Gyani, R. Lakshman Naik, G. Sireesh Reddy

67. ESTIMATION AND OPTIMIZATION OF PROSODIC TO IMPROVE THE QUALITY OF THE ARABIC SYNTHETIC SPEECH 632-639

Abdelkader CHABCHOUB & Adnen CHERIF

68. INTEGRATION OF CONTROL CHARTS AND DATA MINING FOR PROCESS CONTROL AND QUALITY IMPROVEMENT 640-648

E. V. Ramana and P. Ravinder Reddy

69. BINS APPROACH TO IMAGE RETRIEVAL USING STATISTICAL PARAMETERS BASED ON HISTOGRAM PARTITIONING OF R, G, B PLANES 649-659

H. B. Kekre and Kavita Sonawane

70. ADAPTIVE ALGORITHM FOR CALIBRATION OF ARRAY COEFFICIENTS 660-667

K. Ch. Sri Kavya, B. V. Raj Gopala Rao, Gopala Krishna.N, J. Supriyanka, J.V.Suresh, Kota Kumar, Habibulla Khan, Fazal Noor Basha

71. SYNTHESIS AND CHARACTERIZATION OF CATALYSTS CONTAINING MOLYBDENUM AND TUNGSTEN AND THEIR APPLICATION IN PARAFFIN ISOMERIZATION 668-676

Aoudjit Farid

Members of IJAET Fraternity

A-G

STATE OF ART: HAND BIOMETRIC

Sarâh BENZIANE¹ and Abdelkader BENYETTOU²

¹Institute of Maintenance and Industrial Safety, University of OranEs-Sénia, Algeria

²Department of Computer Science, Faculty of Science, University of Science & Technology
Mohamed Boudiaf of Oran, Algeria

ABSTRACT

This paper present a state of art about biometric hand, different techniques used. Biometric is essentially used to avoid risks of password easy to find or Stoll; with as slogan save Time and Attendance. We can note that biometrics is a true alternative to the passwords and other identifiers to make safe the access controls. It makes it possible to check that the user is well the person who it claims to be.

KEYWORDS: Hand, palmprint, geometry, biometric system, identification, authentication, verification.

I. INTRODUCTION

Biometrics is in full growth and tends to join other technologies of safety like the smart card. Within the biometric systems used today, we notice that the hand biometric is one of those, the users most accept because they don't feel persecute in their private life. A survey of 129 users illustrated that the use of hand geometry biometric system at Purdue University's Recreation Centre has many advantages; the survey participants, 93% liked using the technology, 98% liked its ease of use, and specially more no else find the technology intrusive [KUK06].

It's why; nowadays hand biometrics recognition has been developed with a great success for the biometric authentication and identification. The biometric recognition process allows the recognition of a person basing on physical and behavioral features. Because of each person have characteristics which are clean for him: voice, fingerprints, features of his face, his signature... his ADN and by the way hand physionomy and physiology, an overview of such systems can be found in [ROS06]. The hand is the almost appropriate for some situations and scenarios.

For the hand biometric modality, within the main features used; we note: the length and width analysis, the shape of the phalanges, articulations, lines of the hand ...etc

The hand biometrics presents a high ease to use a system based on. Although, the hardware system from time to time makes error incidence's due to the injury of the hand and by the way the hand age. Setting besides that, the systems gives a very high accuracy with a medium security level required. However, for a long term the stability is somehow average and need to be improved. Most of the previous works has elaborated systems based on hand biometric contact [SAN00].

The reminder of this paper is organized as follow. In section 2, we present why we use the hand biometric. In section 3, we describe how does hand biometric system works. In Section 4, we present the hand identification techniques. In Section 5, we present the bottom up feature based methods. In section 6, we present the data capture. In Section 7, we present the hand biometric identification/ authentication. In section 8, we present a tabular representation of the existing method. In last section, we offer our conclusion.

II. WHY HAND BIOMETRIC?

The suitability of a specific biometric to a particular application depends on many issues [50]; amid them, the user acceptability appears to be the most important [JAI97]. For various access control

applications, as immigration, border control and dormitory meal plan access, very distinctive biometrics, e.g., fingerprint and iris, could not be suitable for protecting person's privacy. In such circumstances, it is preferable that the given biometric key be only unique enough for verification but not for identification. The evaluation of a biometric method depends on the reliability, security, performance, cost, user acceptance, life detection, users, size of sensor. One of its advantages is the aging issues, both young and old.

III. HOW DOES HAND BIOMETRIC SYSTEM WORK?

A hand biometric system works like the other systems based on the other modality as fingerprint, voice, iris... Maybe, it can differ only in some few points, like the way to make safe the information. But, generally the scenario bellow (Fig. 1) is used to conceive a hand or another biometric system:

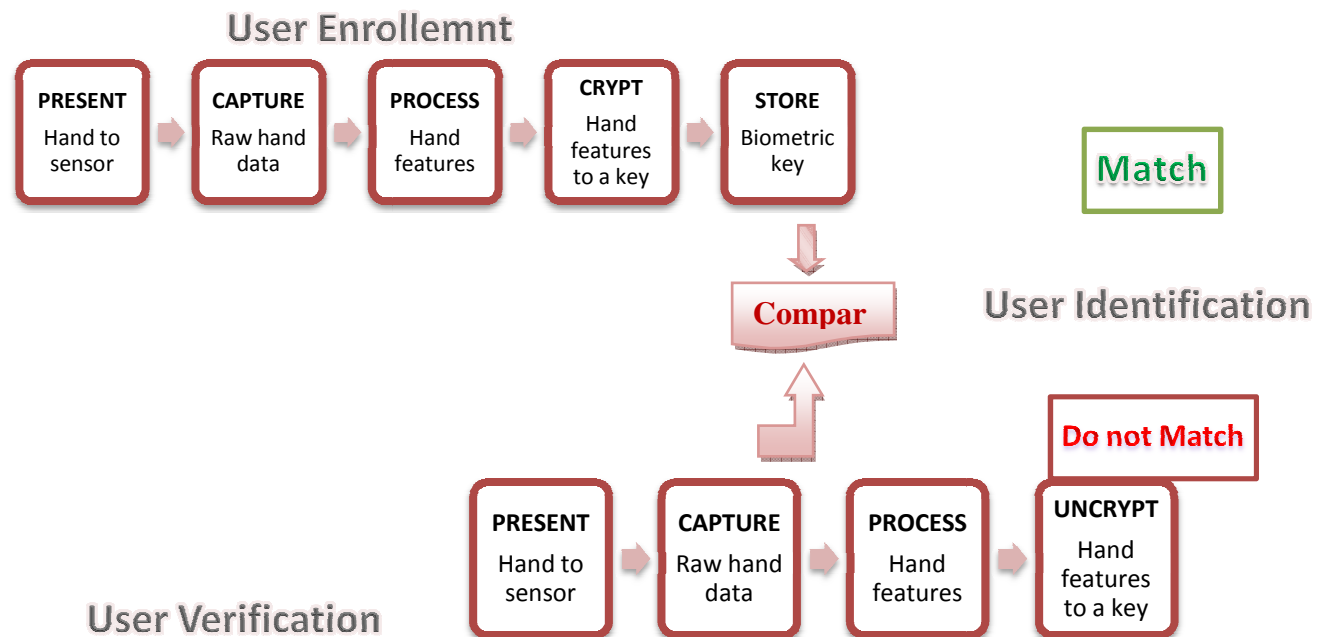


Figure 1 Hand biometric system scenario's

It is based on three basic processes; the enrolment, the verification and the identification. The enrolment phase is used for Adding a biometric identifier to the database. The Verification, more known as one towards one, because it must make sure that the person is whom he/she claim to be by matching against a single record. The Identification, more known to as one against all, since it ought to find who is this individual through a matching against all the records in the database.

IV. HAND IDENTIFICATION

There are three clusters of characteristics which are used in hand identification, which are called, too bottom up features:

- Geometric features; such as the width, length and area of the palm. Geometric features are a rough measurement and they are not sufficiently distinct;
- Line features, principal lines and wrinkles. Line features identify the size, position, depth and length of the various lines and wrinkles on a palm. Although wrinkles are very characteristic and are not easily copied, principal lines may not be satisfactorily distinct to be a reliable identifier;
- Point features or minutiae. Point features or minutiae are similar to fingerprint minutiae and classify, between other features, ridges, ridge endings, bifurcation and dots.

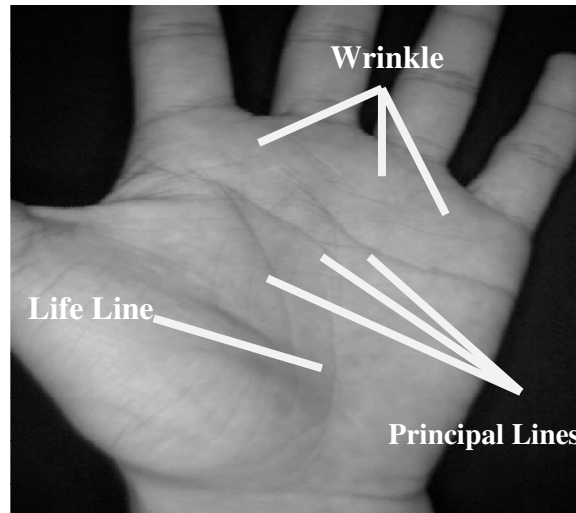


Figure 2 : Hand's Lines

V. BOTTOM-UP FEATURE-BASED METHODS

The human hand is the source of a number of unique physiological characteristics. The main technologies for hand recognition fall into three categories: palmprint technologies – those measuring the unique pattern of prints from the palm of the hand – similar to a fingerprint; Hand geometry measurements – those measuring the shape and size of either all or part of the human hand or fingers; Hand vein patterns – those measuring the distinct vascular patterns of the human hand, including hand dorsum vein and palm vein.

5.1. Palmprint features

They are made up of principal lines, deltapoints, minutiae, wrinkles, singular points and texture, etc...[32]. Several approaches are used for. Within the most popular methods, those considered the palmprint images as textured images which are sole for each person. [9] apply gabor filter for palmprint image analysis using a digital camera where [11] used the wavelets, [16] the Fourier Transform, [44] the local texture energy and [41] the directional line energy features. Therefore, [DUT03] used a set of feature points the length of the major palm lines. Though, in palmprints the creases and ridges often overlies and cross each other. So, [3] has put forward the extraction of local palmprint features by eliminating the creases; but this work is only limited to the extraction of ridges. Where [45] by generating a local gray level directional map; has tried to approximate palmprint crease points.

Generally the steps used for the palmprint based biometric are; first to align and localize palm images by detecting and aligning to inter-finger anchor points: index-middle and ring-pinky junctions. After, to extract with a certain resolution pixel region and down sample on each of the 5 direct multispectral images per hand placement. Then, Process with orthogonal line ordinal filter to generate level I palmprint features. Next, Perform round-robin, single-sample matching of palm features by for example the Maximum Hamming distance over multiple translations [21]. Finally, if the palmprint is used in a multibiometric; so we must fuse the palm print by normalizing the match scores to the same range taking into account the product of the individual match scores.

5.2. Hand geometry features

They are based on the area/size of palm, length and width of fingers. Most of the works in the biometric hand are based on the geometric features [36] [39]; [21] used geometric features and implicit finger polynomial invariants. [SAN00] use user-pegs to constrain the rotation and translation of hand.

5.2.1. Contact hand biometric

Most of the systems proposed and/or used are based on research restricted to significantly old patents and commercial products [26]. These systems are made as the user must push his/her hand on the sensor surface, placing his/her dwells correctly with the guidance's peg. From that process it's possible to extract some features like: length, width and height of the fingers, thickness of the hand, aspect ratio of fingers and palm which make possible the building of small template. Some works based on the systems described above was focused on accuracy. [SAN00][29] have proposed an original and better-off geometric features set and have examined the use of multiple templates for a person using the Gaussian Mixture Models to model each focus. [8] suggested the use of the all contour silhouette of the hand in a straight line for matching.

Although, several studies has shown that the peg-based alignment is not very efficient and can be in some cases the source of failure [SAN00] [26]. So, more recent studies has concentrate their works on a more suitable design of a peg-free system [23] [39][12] [2] [18] [42]. Extraction of the hand from the background is the first step of the processing, to after segment the hand in fingers and palm to get finally the geometric features [39] [12] and the contours related to each one of them [18][42].

5.2.2. Contactless hand biometric

A new approach for hand biometric has been used recently in many work, which is the contactless hand biometric. [20] centered on a peg-free hand recognition, based on EIH-inspired method for robustness against noise.

5.3. Hand vein

To provide fast, accurate and robust personal identification, some authors [34] proposed to use the hand vein as feature identification. [17] gives an overview of the hand-vein application. Current products based vein identification permit single person authentication in less than a second.

5.4. Palmprint & hand geometry features

To mitigate each previous technics problems some authors proposed to use the palmprint and the hand geometry features. Some propose to use two different sensors for each part. In [13], the plamprint and the hand shape are extracted from sensor but the fingerprint is extracted from another sensor. Although, the most interesting is to use as for the other bimodal biometric systems, a single sensor. The most appropriate for this situation is to make use of a digital camera to get only one image to process and with a high resolution. This is what proposed [12] and used; they combined the both features kind with fingerprints information, after examining them and using a simple image acquisition setup.

VI. DATA CAPTURE

We can count three techniques for capturing the hand:

- Off-line, palm prints are inked into paper and after scanned by the palm print system. By the past, the researchers used in their works offline palmprint images and get interesting results [DUT01] [44][SHI01].
- On-line, palm prints are scanned directly as in [44]; which presents survey the use of texture to represent low-resolution palmprint images for online personal identification
- Real-time, palm prints are captured and processed in real-time.

6.1. Resolution quality

The both low and high resolutions are based on some features, and it depends on the application where it's used.

6.1.1. Low resolution

PalmPrint features, which are composed of principal lines, wrinkles, minutiae, delta points, etc., and must quote the features and give some techniques and works for the both. However, features like

principal lines and wrinkles, can be extract from a low resolution image with less than 100 dpi [32][47].

6.1.2. High resolution

For features such as minutiae points, ridges, and singular points, a high image resolution is required for a good extraction with at least 400 dpi (dots per inch) [SHI01].

VII. HAND BIOMETRIC IDENTIFICATION/AUTHENTICATION

7.1. Detection

First for the recognition, we must extract the hand shape from background [28], as well as motions to obtain hand features [SHI01]. Most of the works used can be based on the hand gesture extraction [31]; because the both are using motion information from extracted hand area in sometimes complex background image for the contactless hand biometric.

Some techniques are used like:

- Background subtraction: used mainly for multiple tracking; human (detection in the meeting room), faces and too for the hand detection [22]
- Skin color: Human skin color [10] has been exploited and established to be an efficient feature in many applications from face to hand detection applied in the different color spaces (RGB, HSV, CIE LUV and the CIE LUV). It integrates strong spoof detection and acquisition.

[21]uses the length and the width of the finger.To get the extraction of the hand, when the localization of the hand extremities, the fingertips and the valleys the main problems met are the artifacts and the unsmoothed contour [43].

In some framework, it's both possible to detect a hand and its corresponding shape efficiently and robustly without constraints upon either user or environment. This has long been an area of interest due to its obvious uses in areas such as sign and gesture recognition to name but just two. Boosting is a general method that can be used for improving the accuracy of a given learning algorithm [30].

7.2. Features extraction

7.2.1. Hand geometry

The hand shape integrated acquisition and reduced computational requirements. Several apparatus was issued based on the hand geometry [19] [7][33].

7.2.2. Palmprint

A plamprint pattern is made up of palm lines; principal lines and creases. Line feature matching is known to be strong and present a high accuracy in palmprint verification [32] [47].

Unfortunately, it is difficult to get a high identification rate by means of only principal lines as their similarity amid different people. The texture representation for coarse-level palmprint classification offers a successful technique [44] survey the use of texture to represent low-resolution palmprint images for online personal identification.

We found in [35] that according to the features used for palmprint recognition, we can distinguish within the various palmprint identification techniques three classes: the structural feature based, appearance based and the texture based. For [27], the best palmprint matching approach, in terms of authentication accuracy is those of [35]. This method is based on the comparison of two line-like image areas and the generation one-bit feature code representing at each image location. The success of this method is due to its stability even when the image intensities vary; which were implemented and tested in [27] successfully for the palmprint matching.

When saying Palmprint, we sepakmaily of the major features and the ridges. They reduced need to manipulate hand or pre processing the skin and integrated acquisition. Sometimes, the fingerprints are used because they represent a robust acquisition under adverse conditions.

7.3. Motion

To know more how the articulation of the hand function, studies was done to analyze and synthesize the 3D movement of the hand [4] which was extended to others biometric.

7.4. Translation/rotation

To be used, many problems can be met like the position of the hand in the image, i.e. the way the hand is presented to the sensor mainly in a contactless hand identification situation. [13] employed unconstrained peg-free imaging, based on the efficiencies of the algorithm to achieve illumination, translation and rotation invariant features. Where the acquired images were binarized and in use for feature extraction. The thresholding edge was automatically calculated, by Otsu's approach, once for each acquisition setup.

7.5. Verification

Some works are based on simple classifier as the Mahalanobis distance [21], mean average distance of contours [8]. [5] applied morphological and Sobel edge features to characterize palmprints and used a neural network classifier for their verification. However, this work has shown the utility of inkless palmprint images acquired from the digital scanner instead of the classical way i.e. the acquisition systems using CCD based digital camera [9].

7.6. Virtual interface

Another main approach in the literature implies the 3D surface reconstructing of the hand. [40] has exploited a range sensor to rebuild the dorsal part hand; they used Local shape index values of the fingers. Sometimes to modalize the hand movement, is used a virtual interface [38]

[14] built an exact hand shape using the splines and hand state recovery could be achieved by minimizing the difference between the silhouettes.

The synthesis fingerprint technique can be applied to synthetic palmprint generation.

7.8. Reconstruction

The estimation of hand pose from visual cues is a key problem in the development of intuitive, non intrusive human computer interface. The solution is to recover a 3d Hand pose from a monocular color sequence; using concepts from stochastic visual segmentation, computer graphics and non linear supervised learning. In [24], made contribution in proposing a automatic system that tracks the hand and estimates its 3D configuration in every frame [ATH01], that does not impose any restrictions on the hand shape, does not require manual initialization, and can easily recover from estimation error. It is possible to approach this problem using a combination of vision and statistical learning tools.

VIII. EXISTING METHOD

Different hand biometric (measurement) techniques need different resources from operating systems to enable biometric authentication on the technical basis of measuring a biological characteristic. Next table gives a tabular overview of different features used.

Six features are considered:

Systems	No. of people	No. of sample per person	Pegs	N of template (s)	Feature (s)	Similarity	Performance	Resolution
Zhang [46]	500	6	No	2	Joint palmprint and palm vein verification	Dynamic weight sum	EER 0.0212% and 0.0158%	352 * 288
Ladoux [15]	24	N/N	N/N	N/N	Palm Vein	SIFT	EER 0%.	232x280
Heenaye [6]	200	N/N	N/D	N/D	Dorsal hand vein pattern,	Cholesky decomposition and Lanczos algorithm	FAR 0%, FRR 0%	320x240

Shahin [SAH07]	50	10	N/D	2	Dorsal hand vein pattern	maximum correlation percentage	FAR of 0.02% and FRR of 3.00 %	N/D
Uhl [37]	25	25	No	N/D	Eigen fingers and minutiae Features	Parallel Versus Serial Classifier Combination	97.6% RR at 0.1% FAR)	500 dpi
Zhang [ZHA09]	120	48	No	4	FINGER-KNUCKLE-PRINT	angular distance	FRR 0.01% and FAR 96.83%	N/D
Oden [ODE03]	27	10	No	270	Geometric features and implicit polynomials invariants of fingers	Mahalanobis distance	N/D	N/D

IX. CONCLUSION

In this paper, we considered a state of art of the hand biometric. The hand can be fusion with other biometrics as face fingerprint and many others [25]. The fact that a disgruntled employee or customer or a person with criminal intentions of entitlement of an active employee in her property and thus brings gives unauthorized access, is another security risk that exclude the biometric hand scanners effectively. One of the most important indirect problems of the hand biometric, is the hand geometry imitation. If the person has arthritis, long fingernails, is wearing hand cream or has circulation problems then this will not produce a good reading. The experimental results provide the basis for the further development of a fully automated hand-based security system with high performance in terms of effectiveness, accuracy, robustness, and efficiency. Individual mobility doesn't have a price; hence, Hand Biometric Technologies have to be implemented whenever and wherever possible.

REFERENCES

- [1] Y. Bulatov, S. Jambawalikar, P. Kumar, and S. Sethia. Hand recognition using geometric classifiers. ICBA'04, Hong Kong, China, pages 753–759, July 2004.
- [2] J. Funada, N. Ohta, M. Mizoguchi, T. Temma, K. Nakanishi, A. Murai, T. Sugiuchi, T. Wakabayashi, and Y. Yamada, "Feature extraction method for palmprint considering elimination of creases," Proc. 14th Intl. Conf. Pattern Recognition., vol. 2, pp. 1849 -1854, Aug. 1998.
- [3] G. Gibert, G. Bailly, D. Beutemps, F. Elisei, and R. Brun, "Analysis and synthesis of the 3D movements of the head, face and hand of a speaker using cued speech," Journal of Acoustical Society of America, vol. 118, pp. 1144-1153, 2005.
- [4] C. -C. Han, H.-L. Cheng, C.-L. Lin and K.-C. Fan, "Personal authentication using palmprint features," Pattern Recognition, vol. 36, pp. 371-381, 2003.
- [5] "Feature Extraction of Dorsal Hand Vein Pattern using a fast modified PCA algorithm based on Cholesky decomposition and Lanczos technique", Maleika Heenaye- Mamode Khan, Naushad Mamode Khan and Raja K. Subramanian, World Academy of Science, Engineering and Technology 61 2010
- [6] I. H. Jacoby, A. J. Giordano, and W. H. Fioretti, "Personal identification apparatus," U. S. Patent No. 3648240, 1972.
- [7] A. K. Jain and N. Duta, "Deformable matching of hand shapes for verification," presented at the Int. Conf. Image Processing, Oct. 1999.
- [8] W. K. Kong and D. Zhang, "Palmprint texture analysis based on low-resolution images for personal authentication," Proc. ICPR-2002, Quebec City (Canada).
- [9] Jure Kovac, Peter Peer, Franc Solina (2003), "Human Skin Colour Clustering for Face Detection.", International Conference on Computer as a Tool, pp. 144- 148.
- [10] A. Kumar and H. C. Shen, "Recognition of palmprints using wavelet-based features," Proc. Intl. Conf. Sys., Cybern., SCI-2002, Orlando, Florida, Jul. 2002.

- [11] Y. A. Kumar, D. C.M.Wong, H. C. Shen, and A. K. Jain, "Personal verification using palmprint and hand geometry biometric.", in Proc. 4th Int. Conf. Audio Video-Based Biometric Person Authentication, Guildford, U.K., Jun. 9-C11, 2003, pp. 668-C678
- [12] A. Kumar and D. Zhang, "Personal recognition using shape and texture," IEEE Trans. Image Process, vol. 15,no. 8, pp 2454- 2461, Aug. 2006.
- [13] James J. Kuch and Thomas S. Huang. Vision-based hand modeling and tracking for virtual teleconferencing and telecollaboration. In Proc. of IEEE Int'l Conf. on Computer Vision, pages 666–671, Cambridge, MA, June 1995.
- [14] Pierre-Olivier Ladoux, Christophe Rosenberger, Bernadette Dorizzi, "Palm Vein Verification System based on SIFT matching", Advances in Biometrics : Third International Conference, ICB 2009, Alghero, Italy, June 2-5, 2009 (2009) 1290-1298", DOI : 10.1007/978-3-642-01793-3_130
- [15] W. Li, D. Zhang, and Z. Xu, "Palmprint identification by Fourier transform," Int. J. Patt. Recognit. Art. Intell., vol. 16, no. 4, pp. 417-432, 2002.
- [16] S. Malki, S., Y. Fuqiang, and L. Spaanenburg, "Vein Feature Extraction using DTCNNs", Proceedings 10th IEEE Workshop on CNNA and their Applications (Istanbul, August 2006) pp. 307 – 312.
- [17] Y. L. Ma, F. Pollick, , and W. Hewitt. Using b-spline curves for hand recognition. Proc.of the 17th International Conference on Pattern Recognition (ICPR'04), Vol. 3:274–277, Aug. 2004.
- [18] R. P. Miller, "Finger dimension comparison identification system," U. S. Patent No. 3576538, 1971.
- [19] "Robust hand image processing for biometric application", JugurtaMontalvao, Lucas Molina, JanioCanuto, Pattern Anal Applic (2010) 13:397–407, DOI 10.1007/s10044-010-0185-7
- [20] Oden, A. Ercil, and B. Buke, "Combining implicit polynomials and geometric features for hand recognition," Pattern Recognition Letters, vol. 24, pp. 2145-2152, 2003.
- [21] S. Ribaric, D. Ribaric, and N. Pavesic.Multimodal biometric user-identification system for network-based applications. IEE Proceedings on Vision, Image and Signal Processing, Volume 150, Issue 6:409–416, 15 Dec. 2003.
- [22] Romer Rosales, VassilisAthitsos, Leonid Sigal, and Stan Sclaroff, "3D Hand Pose Reconstruction Using Specialized Mappings", Boston University Computer Science Tech. Report No. 2000-22,Dec. 2000 (revised Apr. 2001), To Appear in Proc. IEEE International Conf. on Computer Vision (ICCV). Canada. Jul. 2001.
- [23] "Information Fusion in Biometrics", Arun Ross and Anil Jain, in Pattern Recognition Letters, Vol. 24, Issue 13, pp. 2115-2125, September, 2003
- [24] A. Jain, A. Ross, and S. Pankanti. A prototype hand geometry-based verification system. Proc. 2nd Int. Conf. on Audio- and video-based personal authentication (AVBPA), Washington, USA, pages 166–171, March 1999.
- [25] "A MULTISPECTRAL WHOLE-HAND BIOMETRIC AUTHENTICATION SYSTEM", Robert K Rowe, UmutUludag, MeltemDemirkus, SulanParthasaradhi, Anil K Jain, IEEE 2007 Biometrics Symposium
- [26] H.Sagawa,M.Takeuchi, "A Method for Recognizing a Sequence of Sign language Words Represented in Japanese Sign Language Sentence",Face and Gesture, pp. 434–439,2000.
- [27] R. Sanchez-Reillo. Hand geometry pattern recognition through gaussian mixture modelling. 15th International Conference on Pattern Recognition (ICPR'00), Volume 2:937–940, 2000.
- [28] Jae-Ho Shin, Jong-Shill Lee, Se-KeeKil, Dong-Fan Shen, Je-Goon Ryu, Eung-Hyuk Lee, Hong-Ki Min, Seung-Hong Hong, "Hand Region Extraction and Gesture Recognition using entropy analysis", IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.2A, PP. 216-223, February 2006
- [29] W. Shu and D. Zhang, "Automated personal identification by palmprint," Opt. Eng., vol. 37, no. 8, pp. 2359-2362, Aug. 1998.
- [30] D. P. Sidlauskas, "3D hand profile identification apparatus," U. S. Patent No . 4736203, 1988
- [31] S. Malki, L. Spaanenburg: Hand Veins Feature Extraction using DT-CNNs Proceedings SPIE 3rd Int. Symposium on Microtechnologies for the New Millennium, Maspalomas, Vol. 6590, 2007-05.

- [32] Z. Sun, T. Tan, Y. Wang, and S.Z. Li, "Ordinal palmprint representation for personal identification", Proc. IEEE Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 279-284, 2005.
- [33] S; Travieso C.M.; Alonso, J.B.; Ferrer M.A.; "Automatic biometric identification system by hand geometry", IEEE 37th Annual 2003 International Carnahan Conference on 14-16, 281 – 284, Oct. 2003.
- [34] Andreas Uhl and Peter Wild, "Parallel Versus Serial Classifier Combination for Multibiometric Hand-based Identification", In M. Tistarelli, M.S. Nixon, editors, Proceedings of the 3rd International Conference on Biometrics 2009 (ICB'09), pp. 950-959, LNCS, 5558, Springer Verlag, 2009
- [35] Vassilis Athitsos and Stan Sclaroff, "3D Hand Pose Estimation by Finding Appearance-Based Matches in a Large Database of Training Views". Technical Report BU-CS-TR-2001-021. A shorter version of this paper is published in the proceedings of IEEE Workshop on Cues in Communication, 2001
- [36] Alexandra L.N. Wong and Pengcheng Shi, "Peg-Free Hand Geometry Recognition Using Hierarchical Geometry and Shape Matching", IAPR Workshop on Machine Vision Applications (MVA02), 2002.
- [37] D. L.Woodard and P. J. Flynn. Personal identification utilizing finger surface features. In CVPR, San Diego, CA, USA, 2005.
- [38] X. Wu, K. Wang, and D. Zhang, "Fuzzy directional energy element based palmprint identification," Proc. ICPR-2002, Quebec City (Canada).
- [39] W. Xiong, C. Xu, and S. H. Ong. Peg-free human hand shape analysis and recognition. Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05), Volume 2:77–80, March 18-23 2005.
- [40] Yoruk, E. Konukoglu, E. Sankur, B. Darbon, J., "Shape-based hand recognition", IEEE transactions on image processing, Vol. 15, Issue 7, Page 1803-1815, 2006
- [41] J. You, W. Li, and D. Zhang, "Hierarchical palmprint identification via multiple feature extraction," Pattern Recognition., vol. 35, pp. 847-859, 2002.
- [42] J. Chen, C. Zhang, and G. Rong, "Palmprint recognition using crease," Proc. Intl. Conf. Image Process., pp. 234-237, Oct. 2001.
- [43] Zhi Liu, Yilong Yin, Hongjun Wang, Shangling Song, Qingli Li, "Finger vein recognition with manifold learning", Journal of Network and Computer Applications 33 (2010) 275–282
- [44] D. Zhang and W. Shu, "Two Novel Characteristics in Palmprint Verification: Datum Point Invariance and Line Feature Matching," Pattern Recognition, vol. 32, no. 4, pp. 691 702, 1999

Authors

Sarâh BENZIANE is assistant professor in computer science; she obtained her magister electronics about mobile robotics. She hold basic degree from computer science engineering. Now, she's working with biometrics system's processing in SMPA laboratory, at the university of Science and Technology of Oran Mohamed Boudiaf (Algeria). She teaches at University of Oran at the Maintenance and Industrial Safety Institute. Her current research interests are in the area of artificial intelligence and image processing, mobile robotics, neural networks, Biometrics, neuro-computing, GIS and system engineering.



Abdelkader Benyettou received the engineering degree in 1982 from the Institute of Telecommunications of Oran and the MSc degree in 1986 from the University of Sciences and Technology of Oran-USTO, Algeria. In 1987, he joined the Computer Sciences Research Center of Nancy, France, where he worked until 1991 on Arabic speech recognition by expert systems (ARABEX) and received the PhD in electrical engineering in 1993 from the USTOran University. From 1988 through 1990, he has been an assistant Professor in the department of Computer Sciences, MetzUniversity, and Nancy-I University. He is actually professor at USTOran University since 2003. He is currently a researcher director of the Signal-Speech-Image- SIMPA Laboratory, department of Computer Sciences, Faculty of sciences, USTOran, since 2002. His current research interests are in the area of speech and image processing, automatic speech recognition, neural networks, artificial immune systems, genetic algorithms, neuro-computing, machine learning, neuro-fuzzy logic, handwriting recognition, electronic/electrical engineering, signal and system engineering.



USING DYNAMIC DUAL KEYS ENCRYPTION ALGORITHM AS PARTIAL ENCRYPTION FOR A REAL-TIME DIGITAL VIDEO

Abdul Monem S. Rahma¹ and Basima Z.Yacob²

¹Computer Science Department, University of Technology, Baghdad, Iraq

²Computer Science Department, University of Duhok, Duhok, Kurdistan Iraq

ABSTRACT

Advances in digital video transmission have increased in the past few years. Security and privacy issues of the transmitted data have become an important concern in multimedia technology. Digital video stream is quite different from traditional textual data because interframe dependencies exist in digital video. Special digital video encryption algorithms are required because of their special characteristics, such as coding structure, large amount of data and real-time constraints. This paper presents a real-time partial encryption to digital video technique depends on Dynamic Dual Key Encryption Algorithm Based on joint Galois Fields which is fast enough to meet the real-time requirements with high level of security. In this technique the I-frame (Intra-frame) of the digital video scene is extracted and decomposed the color picture into its three color channels: luma channel (Y) and two chrominance channels Cb and Cr, with note that the frames of digital video is in YCbCr color system, the Dynamic Dual Key Encryption Algorithm Based on joint Galois Fields is applied to the Y channel. The encryption technique achieves best timing results, and it provides high level of security by its great resistant against brute force attacks.

KEYWORDS: Encryption digital video, partial encryption for Digital video, Digital video encryption in real time.

I. INTRODUCTION

In the digital world nowadays, the security of digital images/videos becomes more and more important since the communications of digital products over network occur more and more frequently. In addition, special and reliable security in storage and transmission of digital images/videos is needed in many digital applications, such as pay-TV, broadcasting, confidential video conferencing and medical imaging systems, etc. Normal data, such as program code or text, has much less redundancy in its structure. These factors make providing secure digital video a challenge. Various encryption algorithms have been proposed in recent years as possible solutions for the protection of the video data. Large volume of the video data makes the encryption difficult using traditional encryption algorithms. Often, we need the encryption to be done in real-time. The naïve approach for video encryption is to treat video data as text and encrypt it using standard encryption algorithms like AES (Advanced Encryption Standard) or DES (Data Encryption Standard). The basic problem with these encryption algorithms is that they have high encryption time making them unsuitable for real-time applications like PAY-TV, Pay-Per View and Video On Demand (VOD) etc. A unique characteristic of video data is that, even though information rate is very high, information value is very low.

This paper presents an efficient partial encryption technique depends on Dynamic Dual Key Encryption algorithm Based on joint Galois Fields for real-time video transmission.

The Dynamic Dual Key Encryption algorithm Based on joint Galois Fields is considered as a stream of bits and the technique uses dual key, first key (control key) to determine the length of bits block and the second one is used for encryption according to the equation that used addition and multiplication based on mathematical theory of Galois field $GF(2^n)$. Each block (3, 4, 5, or 6) bits size in this algorithm are interpreted as finite field elements using a representation in which a 3, 4, 5 or 6 bits with bits $b_0 b_1 b_2$, $b_0 b_1 b_2 b_3$, $b_0 b_1 b_2 b_3 b_4$ or $b_0 b_1 b_2 b_3 b_4 b_5$ represents the polynomial consecutively, this algorithm is existing and introduced in details in [1].

We apply the encryption algorithm to a part of I-frames of video, exclusively on Y Channel of YCbCr color vector. This technique is fast enough to meet the real-time requirements, in addition it provides high level of security by its great resistant against brute force attacks. To decrypt the ciphertext with 128 bits, the attacker needs $1.86285884e + 204$ of possibilities of keys as minimum and $1.80032832e + 399$ as maximum [1].

The paper is organized as follows. Section 2 presents related work, Section 3 introduces digital video preliminaries. Section 4 and 5 present the methodology of partial encryption and decryption algorithm video consecutively. In Section 6 the suggested technique for partial video encryption is presented. Section 7 shows the experimental results for proposed technique of partial video encryption, Discussion the proposed technique for partial video encryption is presented in section 8. Finally, conclusions are provided in Section 9.

II. RELATED WORK

Many video encryption algorithms have been proposed which encrypt only selected parts of the data. Meyer and Gadget [2] have designed an encryption algorithm named SECmpeg which incorporates selective encryption and additional header information. In this encryption selected parts of the video data like Headers information, I-blocks in P and B frames are encrypted based on the security requirements. Qiao and Nahrstedt [3] proposed a special encryption algorithm named video encryption algorithm in which one half of the bit stream is XORed with the other half. The other half is then encrypted by standard encryption algorithm (DES). The speed of this algorithm is roughly twice the speed of naive algorithm, but that is arguably still the large amount of computation for high quality real-time video applications that have high bit rates [4]. Some of the other encryption algorithms are based on scrambling the DCT coefficients. Tang's [5] scrambling method is based on embedding the encryption into the MPEG compression process. The basic idea is to use a random permutation list to replace the zig-zag order of the DCT coefficients of a block to a 1×64 vector. Zeng and Lie [6] extended Tang permutation range from block to segment, with each segment consisting of several macroblocks. Within each segment, DCT coefficients of the same frequency band are randomly shuffled within the same band. Chen, et. al [7] further modified this idea by extending the permutation range from a segment to a frame. Within a frame, DCT coefficients are divided into 64 groups according to their positions in 8×8 size blocks, and then scrambled inside each group. Apart from shuffling of the I frames, they also permuted the motion vectors of P and B frames. In order to meet the real-time requirements, Shi, et. al [8] proposed a light-weight encryption algorithm named Video Encryption Algorithm (VEA). It uses simple XOR of sign bits of the DCT coefficients of an I frame using a secret m-bit binary key. The algorithm was extended as Modified Video Encryption Algorithm (MVEA) [9] wherein motion vectors of P and B frames are also encrypted along with I frames.

III. DIGITAL VIDEO PRELIMINARIES

Digital video consists of a stream of images captured at regular time intervals, where the digital image is a discrete two-dimensional function, $f(x, y)$ which has been quantized over its domain and range [10]. Without loss of generality, it will be assumed that the image is rectangular, consisting of Y rows and X columns. The resolution of such an image is written as $X \times Y$. Each distinct coordinate in an image is called a pixel color space and each color pixel is a vector of color components.

The Color spaces provide a standard method of defining and representing colors. Each color space is optimized for a well-defined application area [11]. The most popular color models are RGB (used in computer graphics); and YCrCb (used in video systems). Processing an image in the RGB color

space, with a set of RGB values for each pixel is not the most efficient method. To speed up some processing steps many broadcast, video and imaging standards use luminance and color difference video signals, such as YCrCb, this color space is widely used for digital video. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr).

In the RGB representation the channels are very correlated, as all of them include a representation of brightness, in which the brightness information can be recognized from R, G and B channels shown separately. But in YCbCr representation the luminance information of (Y) component is more than chrominance information of (Cb and Cr) components [12].

A color in the RGB color space is converted to the YCrCb color space using the following equation[13]:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \dots\dots\dots(1)$$

While the inverse conversion can be carried out using the following equation:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0.000 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0.000 \end{bmatrix} \begin{bmatrix} Y - 16 \\ C_b - 128 \\ C_r - 128 \end{bmatrix} \dots\dots\dots(2)$$

Digital video stream is organized as a hierarchy of layers called: Sequence, Group of Pictures (GOP), Picture, Slice, Macroblock and Block. The Sequence Layer consists of a sequence of pictures organized into groups called GOPs. Each GOP is a series of I, P and B pictures [14]. I pictures are intraframe coded without any reference to other pictures. P pictures are predictively coded using a previous I or P picture. B pictures are bidirectionally interpolated from both the previous and following I and/or P pictures [7].

Each picture is segmented into slices, where a picture can contain one or more slices. Each slice contains a sequence of macroblocks where a macroblock consists of four luminance blocks (Y) and two chrominance blocks (Cb and Cr). Each block is organized into a matrix of 8x8 pixel samples with a macroblock covering a 16 x 16 pixel area, Figure (1) shows the Structural hierarchy of digital video.

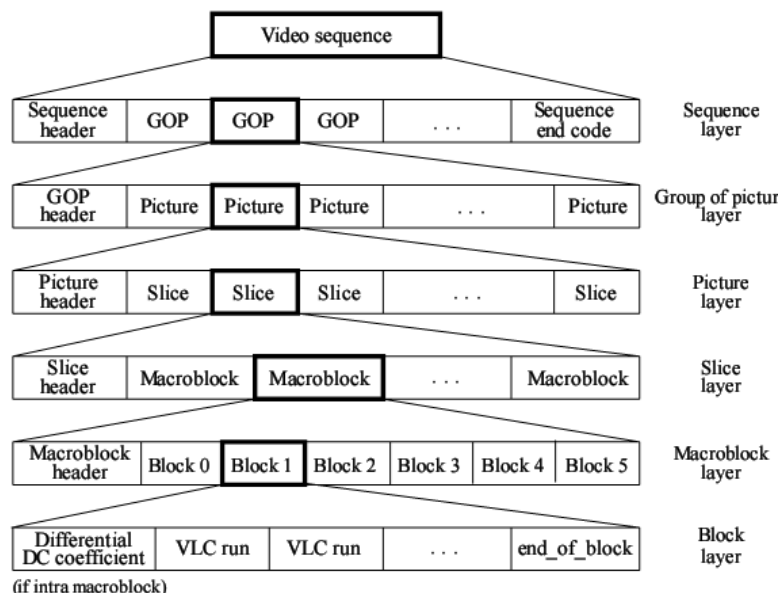


Figure 1: Structural hierarchy of Digital video.

The properties of the I, P, and B frames can help further improve the encryption and decryption performance. Since B frames depend on I or P frames, and P frames depend on the closest preceding I frame, we need only encrypt the I frames while leaving the P and B frames untouched. Without I frames, one cannot decode P and B frames.

For decryption the same steps of encryption are applied but with reverse equation's operations are performed [1].

VI. PARTIAL VIDEO ENCRYPTION TECHNIQUE

The Suggested Technique model consists of two parts; the main stages of the first part are started from reading video file (with note that the frames of digital video is in YCbCr color system), converting it into frames, the output of this stage is frames in YCbCr color representation, the last stage deals with selecting the I-frame. In the second part of system, the Dynamic Dual Key Encryption Algorithm Based on joint Galois Fields is applied on Y-channel of I-frame, then reconstructing the video file before broadcasting. At the receiver side, the video file will be converted into frames, and applying the Dynamic Dual Key decryption Algorithm Based on joint Galois Fields on Y-channel of I-frame, Figure 2 illustrates the steps of proposed system.

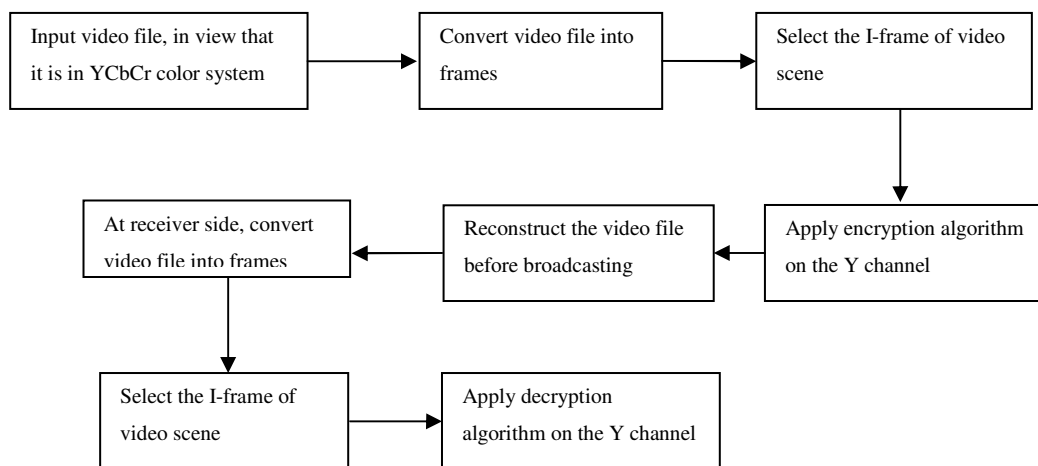


Figure (2): The steps of partial encryption Technique

VII. EXPERIMENTAL RESULTS

Advanced Encryption Standard (AES) is an algorithm of the first category which is used nowadays in communication and encrypted video broadcasting, and it provides much higher security level than DES and perform it in 3 to 10 less computational power than 3-DES [15], it has better performance than DES, 3DES, and RC2 [16], based on these facts, AES is to be compared with proposed technique. The following tables represent the experimental results for the speed of the partial video encryption based on Dynamic Dual Key Encryption algorithm, and AES algorithm.

Table 1: The encryption and decryption times for AES algorithm using key size 128 bit on I-frame

Security Algorithm	I-Frame Name	Size of Frame KB	Encryption time (Second)	Decryption time (Second)
AES-Rijndael	Car	60	8	12
	Wedding	1180	175	260
	xylophone	225	28	46

Table 2: The encryption and decryption times for Dynamic Dual Key Encryption algorithm on I-frame.

Security Algorithm	I-Frame Name	Size of Frame KB	Encryption time (Second)	Decryption time (Second)
Dynamic Dual algorithm	Car	60	0.656	1.282
	Wedding	1180	12.468	28.594
	xylophone	225	2.312	5.438

From Tables 1 and 2, Can be observed that the Dynamic Dual encryption algorithm is approximately 13 times faster than AES encryption and 9 times faster than AES decryption.

The sample test video sequences include videos like Car, Wedding, and xylophone. Some of the test videos along with their frame numbers are shown in Figure 3, Figure 4 and Figure 5.

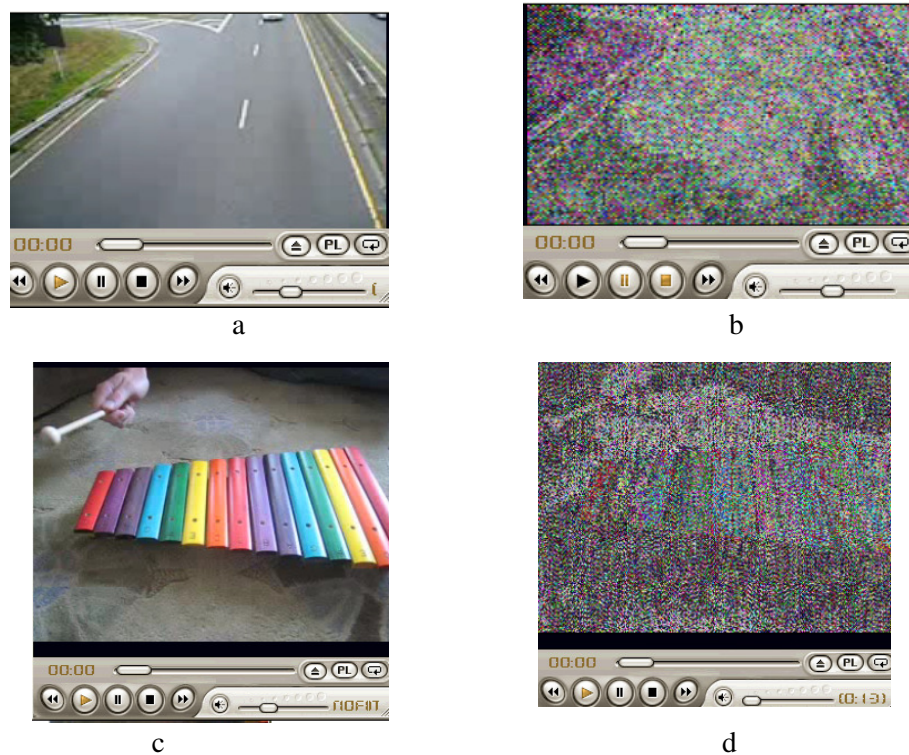


Figure (3): The encryption results after applying partial encryption based on Dynamic Dual keys algorithm for the 1st frame in “Car” and xylophone video. a)Original I-frame of car video b) car I-frame after encryption c) Original I-frame of xylophone video d) xylophone I-frame after encryption.

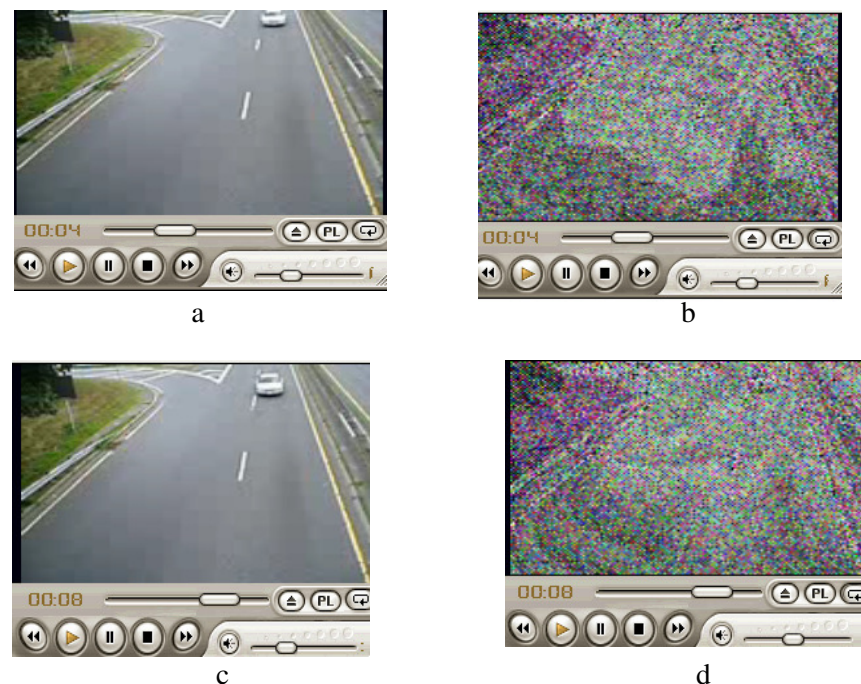


Figure (4): The effect of the partial encryption based on Dynamic Dual keys algorithm on Car Video Frames is used as test object. (a)Original car film after 4 seconds (b) Encryption car film after 4 seconds (c) Original car film after 8 seconds (d) Encryption car film after 8 seconds

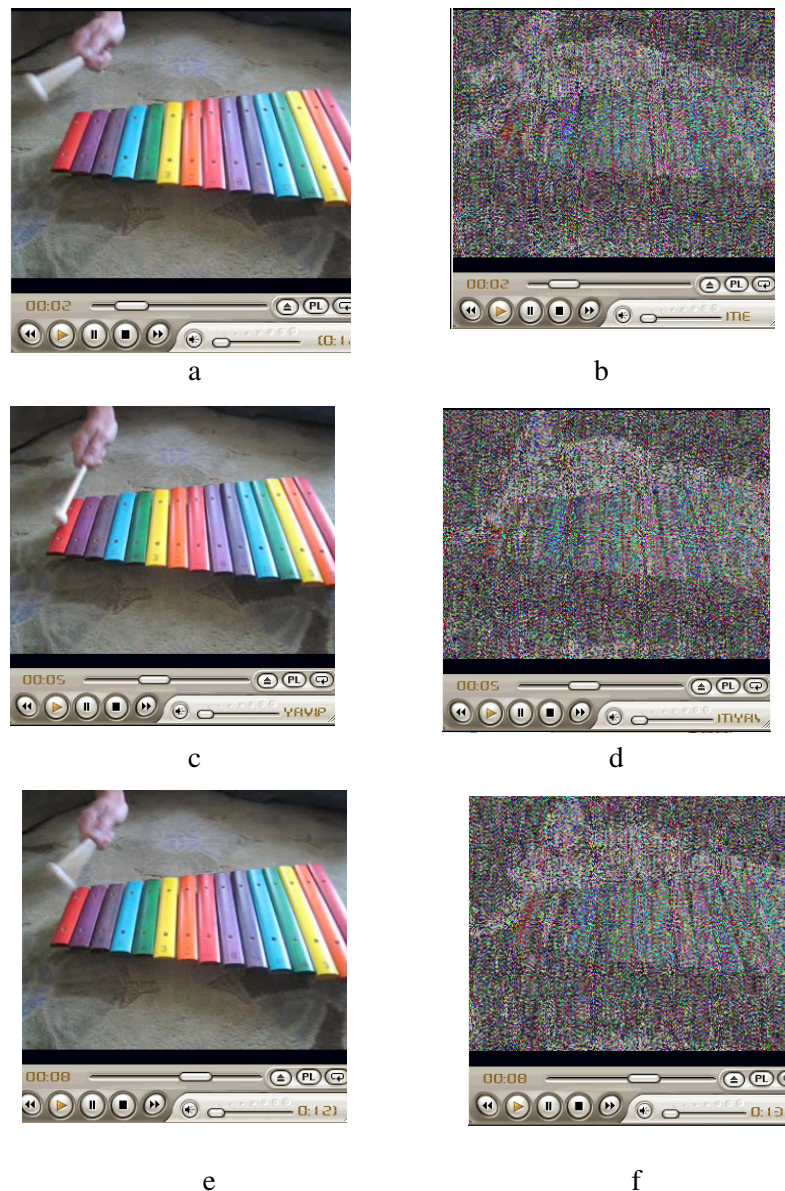


Figure (5) The effect of the partial encryption based on Dynamic Dual keys algorithm on xylophone Video Frames is used as test object (a) Original xylophone film after 2 seconds (b) Encryption xylophone film after 2 seconds (c) Original xylophone film after 5 seconds (d) Encryption xylophone film after 5 seconds (e) Original xylophone film after 8 seconds (f) Encryption xylophone film after 8seconds.

The designed technique and AES algorithm both has been implemented successfully using visual basic 6 Programming language and also implemented with processor of Pentium III (3.40 GHZ) and 3GB of RAM on windows XP.

VIII. DISCUSSION

The comparison between the speed of partial encryption digital video technique which uses AES algorithm, and the partial encryption digital video technique which uses Dynamic Dual Keys algorithm ,it can be seen from tables 1 and 2 that the speed of the technique which uses Dynamic Dual Key Encryption algorithm is faster than the technique which uses AES algorithm, whereas the technique which uses Dynamic Dual Keys algorithm gets the best results , and it is approximately 13 time faster than AES encryption and 9 times faster than AES decryption. Because of the high security obtained by Dynamic Dual Key Encryption algorithm this will make the proposed technique high security.

No variation has been made in the digital video structure by the proposed technique, because of making use of the present broadcasting technique; whereas a change has been made in the part of complete structure.

IX. CONCLUSION

In this paper, we have proposed a new partial digital video encryption technique. The proposed technique which encrypts only Y-channel from I-frame of the digital video scene will reduce the encryption and decryption time, in addition to its high security depending on Dynamic Dual Key Encryption algorithm which uses dynamic block cipher and dual keys. All these properties will make the proposed technique suitable for Real-Time Application RTA.

REFERENCES

- [1] Abdul Monem S. Rahma , and Basima Z.Yacob "The Dynamic Dual Key Encryption Algorithm Based on joint Galois Fields", International Journal of Computer Science and Network Security, VOL.11 No.8, August 2011.
- [2] J. Meyer and F. Gaegast, "Security Mechanisms for Multimedia Data with the Example MPEG-1 Video", Project Description of SEC MPEG, Technical University of Berlin, Germany, May 1995.
- [3] L. Qiao and Klara Nahrstedt, "A New Algorithm for MPEG Video Encryption", In Proc. of First International Conference on Imaging Science System and Technology, pp 21–29, 1997.
- [4] Borko Furht and Darko Kirovski, "Multimedia Encryption Techniques, Multimedia Security Handbook," CRC Press LLC ,Dec. 2004 .
- [5] L.Tang, "Methods for Encrypting and Decrypting MPEG Video Data Efficiently", In Proc. of ACM Multimedia, Boston, pp 219-229, 1996.
- [6] W. Zeng and Sh. Lei, "Efficient Frequency Domain Selective Scrambling of Digital Video", In Proc. of the IEEE Transactions on Multimedia, pp 118-129, 2002.
- [7] Z. Chen, Z. Xiong, and L. Tang, "A novel scrambling scheme for digital video encryption". In Proc. of Pacific-Rim Symposium on Image and Video Technology (PSIVT), pp 997–1006, 2006.
- [8] C. Shi, S. Wang, and B. Bhargava, "MPEG Video Encryption in Real time Using Secret Key Cryptography", In Proc. of International Conference on Parallel and Distributed Processing Techniques and Applications, Las Vegas, NV, 1999.
- [9] C. Shi and B. Bhargava, "A Fast MPEG Video Encryption Algorithm", In Proc. of ACM Multimedia, Bristol, UK, pp 81-88, 1998.
- [10] Robert M. Gray and David L. Neuhoff. "Quantization", IEEE Transactions on Information Theory, 44(6):1.63, October 1998.
- [11] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", Second Edition, Printice Hall Inc, 2002.
- [12] Iain E. G. Richardson. "H.264 and MPEG-4 Video Compression" The Robert Gordon University, Aberdeen, John Wiley & Sons Ltd, UK, 2003.
- [13] Li & Drew, "Fundamentals of Multimedia ", Chapter 5, Prentice Hall 2003.
- [14] P. N. Tudor. "MPEG-2 video compression". In Electronics and Communication Engineering Journal, December - 1995.
- [15] J. Dray, "Report on the NIST Java AES Candidate Algorithm Analysis", NIST ,1999.
- [16] D. S. Abd Elminaam, H. M. Abdual Kader, and M. M. Hadhoud, "Evaluating The Performance of Symmetric Encryption Algorithms", International Journal of Network Security, Vol.10, No.3, PP.216–222, May 2010.

Authors

Abdul Monem Saleh Rahma awarded his MSc from Brunel University and his PhD from Loughborough University of technology United Kingdom in 1982, 1985 respectively. He taught at Baghdad university department of computer science and the Military Collage of Engineering, computer engineering department from 1986 till 2003. He fills the position of Dean Asst. of the scientific affairs and works as a professor at the University of Technology Computer Science Department .He published 82 Papers in the field of computer science and supervised 24 PhD and 57 MSc students. His research interests include Cryptography, Computer Security, Biometrics, image processing, and Computer



graphics. And he Attended and Submitted in many Scientific Global Conferences in Iraq and Many other countries.

Basima Zrkqo Yacob received the B.Sc. degree in Computer Science from Mosul University, Iraq, in 1991, The MSc. Degree in computer science from University of Duhok ,Iraq in 2005. Currently she is a PhD student at faculty of computer science at Duhok University



DESIGN AND PROTOTYPING OF A MINIATURIZED SENSOR FOR NON-INVASIVE MONITORING OF OXYGEN SATURATION IN BLOOD

Roberto Marani, Gennaro Gelao and Anna Gina Perri

Electrical and Electronic Department, Polytechnic of Bari, via E. Orabona 4, Bari - Italy

ABSTRACT

In this paper a new sensor for the non-invasive monitoring of the oxygen saturation in blood has been designed, realized and tested to obtain an ultra-small device having very high noise immunity. This goal has been reached by using a particular integrated circuit, the PsoC (Programmable System on Chip), which integrates two programmable arrays, one analogue and one digital, to obtain a device with very large capabilities. We have configured the PsoC and developed the electronic interfaces. The proposed design allows the acquisition of the continuous component of the signal and the data elaboration has been done in place using a local CPU, without requiring to pass data to an external computer.

KEYWORDS: *Oxyhaemoglobin Saturation, Spectrophotometric Method, Pulse Oximeters, Electronic Interfaces and Data Processing, Sensor Prototyping and Testing.*

I. INTRODUCTION

Non-invasive health monitoring is the main goal of modern electronic applications to medicine. In particular, among the most critical vital parameters, one can find the oxygen saturation of oxyhaemoglobin HbO₂. Currently the standard procedure for monitoring gases in blood is to take arterial blood samples. This is time consuming for the nurses and stressful particularly for those patients with cardiac respiratory or renal insufficiency, i.e. requiring a continuous monitoring.

Several invasive methods for continuous monitoring have been proposed, based on the use of catheter or optical fibre sensors, but they have many problems such as inevitable pain for the patient, possible infections, long term drift caused by blood's substances deposition on the catheter, need for hospitalization, and last but not least, the high cost.

In order to overcome these problems, there is an effort to develop other devices with better characteristics, which allow mainly the non-invasive, continuous monitoring with good accuracy. Among these devices, the pulse oximeter, which senses the oxygen saturation in blood using non-invasive optical sensors, seems to be the best [1]. Although this device is typically used in hospitals, it still has some drawbacks that should be solved in order to make this device available even for home proposes without the assistance of registered nurses. Moreover, among the required enhancements, it has to be cheap, small, user-friendly, accurate and noise immune.

In this paper we present a new pulse oximeter, which has been realized and tested at the Electronic Devices Laboratory (Electrical and Electronic Department) of Polytechnic of Bari.

The proposed sensor, designed in order to obtain a cheap device with reduced size and very high noise immunity, uses a single chip, such as the PSoC (Programmable System on Chip) [2], produced by Cypress MicroSystems, which, through its programmability, i.e. its capability within hardware and software to change, allows the signal acquisition and conditioning of the whole system on a single chip.

In Section 2 we have described the front end of the proposed pulse oximeter, while in Section 3 the obtained results are analyzed and discussed. The conclusions and future scope are illustrated in Section 4.

II. FRONT END OF THE PROPOSED PULSE OXIMETER

The pulse oximetry is a spectrophotometric method for non-invasive measurement of the arterial oxygen saturation, SpO_2 , and pulse [3]. It is based on the different light-absorbing characteristics of oxyhaemoglobin (HbO_2) and deoxyhemoglobin (Hb) at two different wavelengths, typically 660 nm (RED) and 940 nm (IR), and on the pulsatile nature of arterial blood flow [4]. Of course, the optical sensor measurements of HbO_2 and Hb are dependent on the concentration of these molecules in blood. With pulse oximeters, a finger or earlobe probe is used. A red light-emitting diode (LED) and an infrared LED is located, as sources, on one side of the probe, and a couple of photodiodes, as receivers, are located on the other side.

The method relies on difference in the absorption spectra of oxygenated and de-oxygenated hemoglobin. The ratio between these, as shown in [3], has a peak at approximately 660 nm and at higher wavelengths the ratio is lower than one.

Conventionally the previous two wavelengths are used since the absorption ratio is large and small at those wavelengths, respectively. This minimizes the uncertainty of the SpO_2 measurement. The measured absorption is then displayed as an estimate of the arterial oxygen saturation, SpO_2 . The sensor, applied to a finger or to earlobe, can work on the transmitted light, realizing in this way a transmission pulse oximeter, or on the reflected light, as a reflectance pulse oximeter [5-9]. The equipment is electrically the same in both cases.

Fig. 1 shows the block scheme of the proposed pulse oximeter.

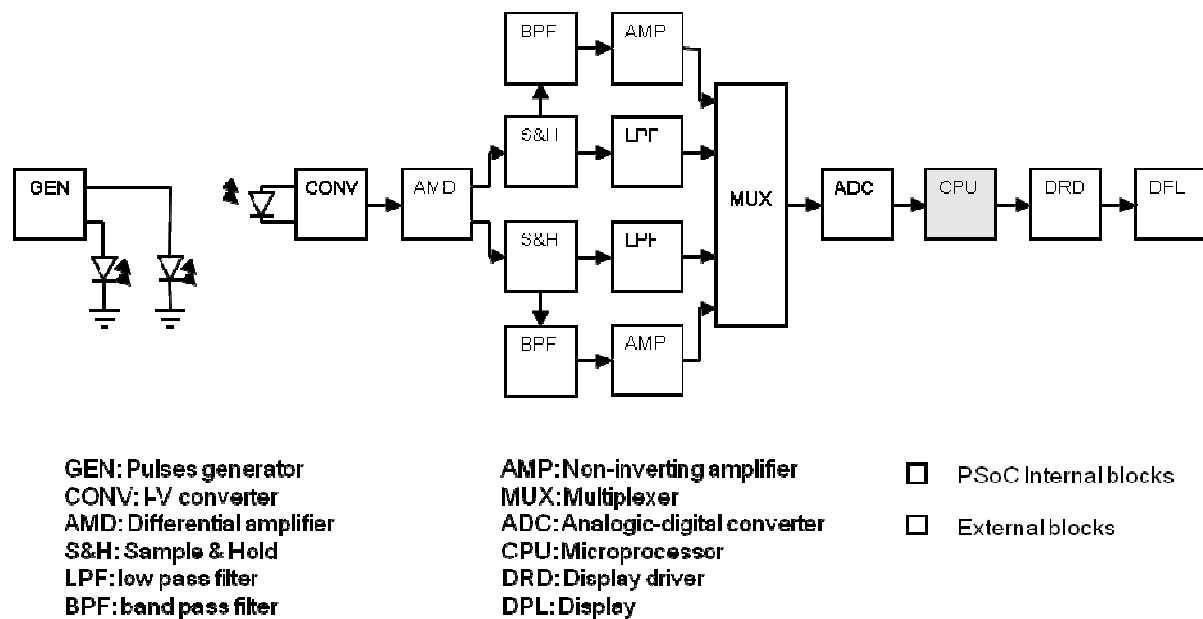


Figure 1. Block scheme of the proposed pulse oximeter.

We have used the PSoC CY8C27443 produced by Cypress MicroSystems [2], a family of devices which allows the implementation of systems on a single chip that contains both analogue and digital programmable blocks, thus allowing the synergic management of analogue and digital signals in a single programmable device, reducing in this way the number of integrated circuits on board.

LEDs are powered by a sequence of square pulses, as shown in Fig. 2, 0.2 ms long at a frequency of 500 Hz and with a phase difference of 1 ms, obtained by an internal PSoC oscillator.

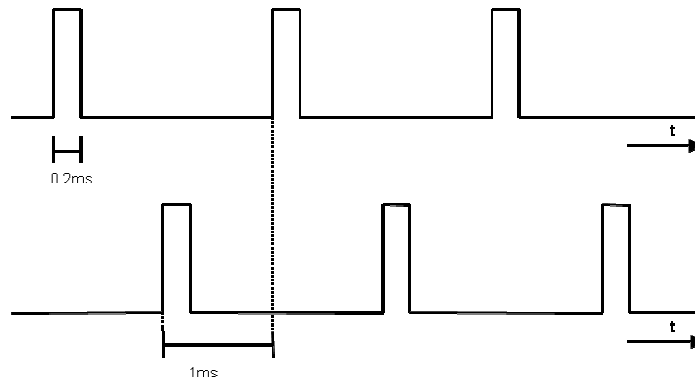


Figure 2. Control signal of the power on the LED.

For each LED we have placed a photodiode on the other side of the finger to collect the transmitted light (Fig. 3). This layout allows us to have a larger collection efficiency.

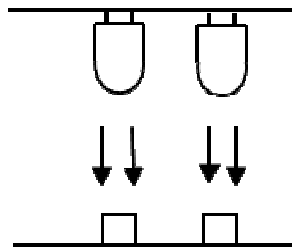


Figure 3. Sensor's scheme.

The light signal measured by the photodiode can be divided in two components: A and B [3]. Component A is the light signal during a systole and is a function of the pulsations of oxygenated arterial blood, while component B is the light signal (during a diastole) that has a constant intensity and is a function of various tissues (i.e. skin pigment, muscle, fat, bone, and venous blood). The pulse oximeter divides the pulsatile absorption of component A by the background light absorption of component B, at the two different wavelengths (RED and IR), to obtain an absorption ratio, R:

$$R = (A_R/B_R) / (A_{IR}/B_{IR}) \quad (1)$$

The photodiode transforms the light signal into an electrical signal that is amplified and converted into digital information.

The current generated by the photodiode is the input of a current-voltage converter working in differential mode, followed by INA105 amplifier (Fig. 4), used to obtain the signal in single ended mode.

The resulting amplifier topology is then that of an instrumental amplifier but with inputs placed in at different nodes, allowing in this way an high noise immunity at the input since most of the noise is a common mode noise.

After the amplifier, the acquisition system splits in two just alike channels, each of them obtained using a Sample & Hold, synchronized with the pulse generator that feeds the LEDs. In this way, it is possible to distinguish between the signal corresponding to red wavelength and the one corresponding to infrared wavelength.

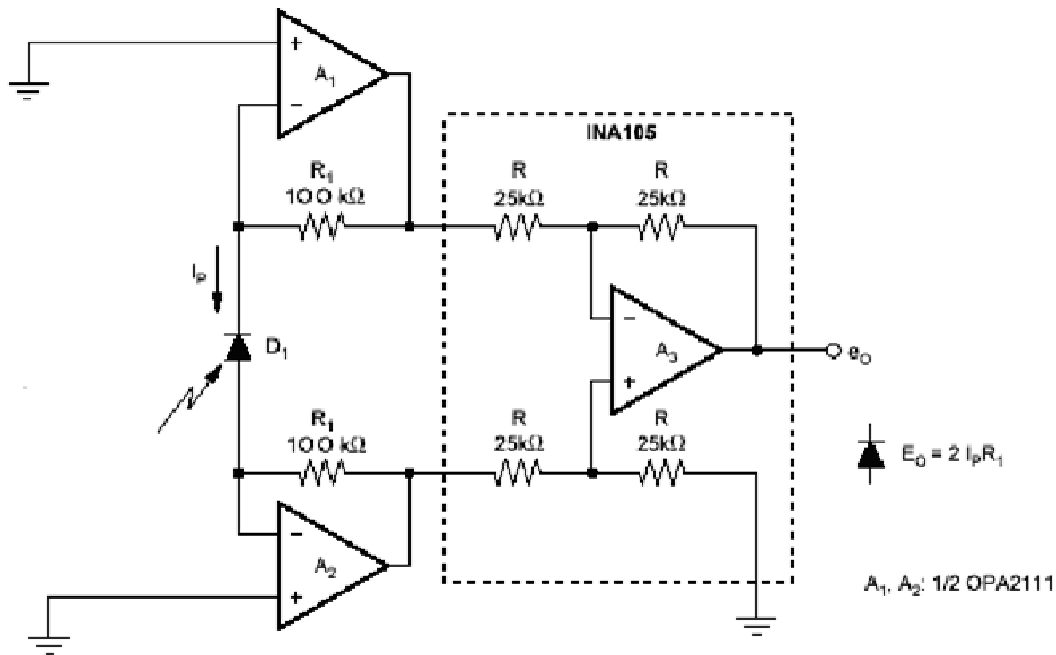


Figure 4. Converter and amplifier circuit configuration.

The oxygen partial pressure in blood has a power spectra concentrated around 1 Hz, while our signal has also high frequency noise, coming from our sampling, and a low frequency noise, down to the continuum, coming from light background, photodiode and from amplifier ($1/f$ noise). For this reason we have applied a band pass (BP) filter realized with two second order stages, both in Sallen-Key configuration [10]. First stage is a low pass filter, the second a high pass filter.

Subsequently the signal goes to a non-inverting amplifier, which insulates the filter from the Analog-Digital Converter (ADC) load, and drives the ADC input.

Based on the red/infrared absorption ratio, defined by the Eqn. (1), an algorithm within the pulse oximeter allows to measure SpO_2 , as a function of the measured magnitude at the systolic and diastolic states on the two photoplethysmograms:

$$SpO_2 = (S_R/B_R) / (S_{IR}/B_{IR}) \quad (2)$$

where S_R is the peak-to-peak red photodiode signal and B_R is the red photodiode continuous component, measured at the systolic and diastolic states respectively, and likewise for S_{IR} and B_{IR} .

Since we need also the continuous components, we have used PSoC's internal component to create a low pass filter with a cut frequency at 200 mHz.

To digitalize the signal, we used a 12 bit ADC available inside the PSoC, and, since we had only one ADC, we had to multiplex the signals to the ADC input under software control.

The digitized signal is then passed to the PSoC's CPU where both signals are normalized and from these data the partial pressure of SpO_2 is computed and shown on a 3-digit display.

With regard to the LEDs, they have a tight band emission spectra with peaks concentrated at 660 nm (R) and 940 nm (IR), as shown in Fig. 5.

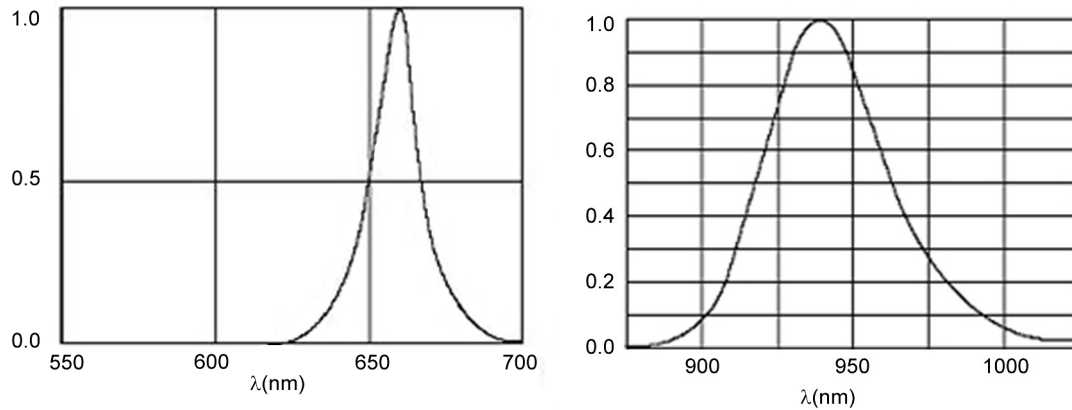


Figure 5. Emission spectra of the two LEDs normalized at their maximum value.

Furthermore their emission is concentrated in a narrow cone in the forward direction.

LEDs are driven with a 50 mA current for the red light and 60 mA for the infrared light.

The photodiode has a wide acceptance angle to maximize the light collection so that it could efficiently collect all the light diffused inside the finger. Its spectral sensibility has a peak at 850 nm, as shown in Fig. 6, and it is at least 60% (compared to the peak) in the bands between 600 nm and 1000 nm, with a good covering of the emission bands of LEDs.

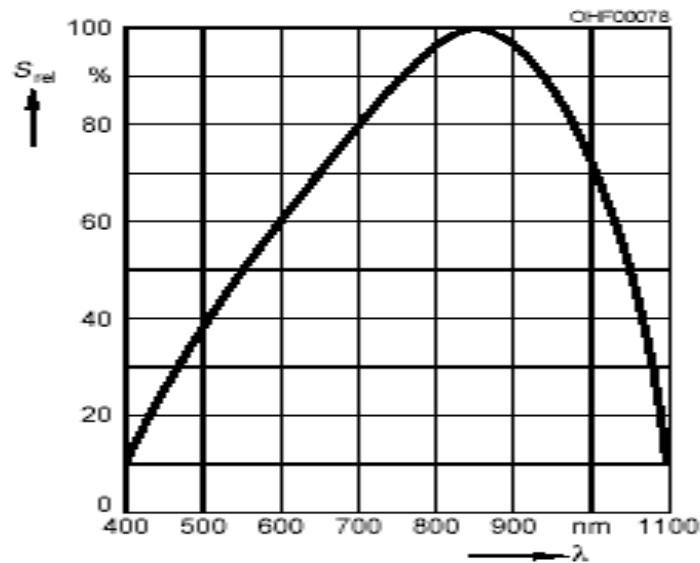


Figure 6. Relative spectral sensibility.

III. ANALYSIS OF RESULTS AND DISCUSSION

With the proposed design, it has been possible to obtain a gain, whose value is about unitary between 0.66 Hz and 3.2 Hz, as shown in Fig. 7.

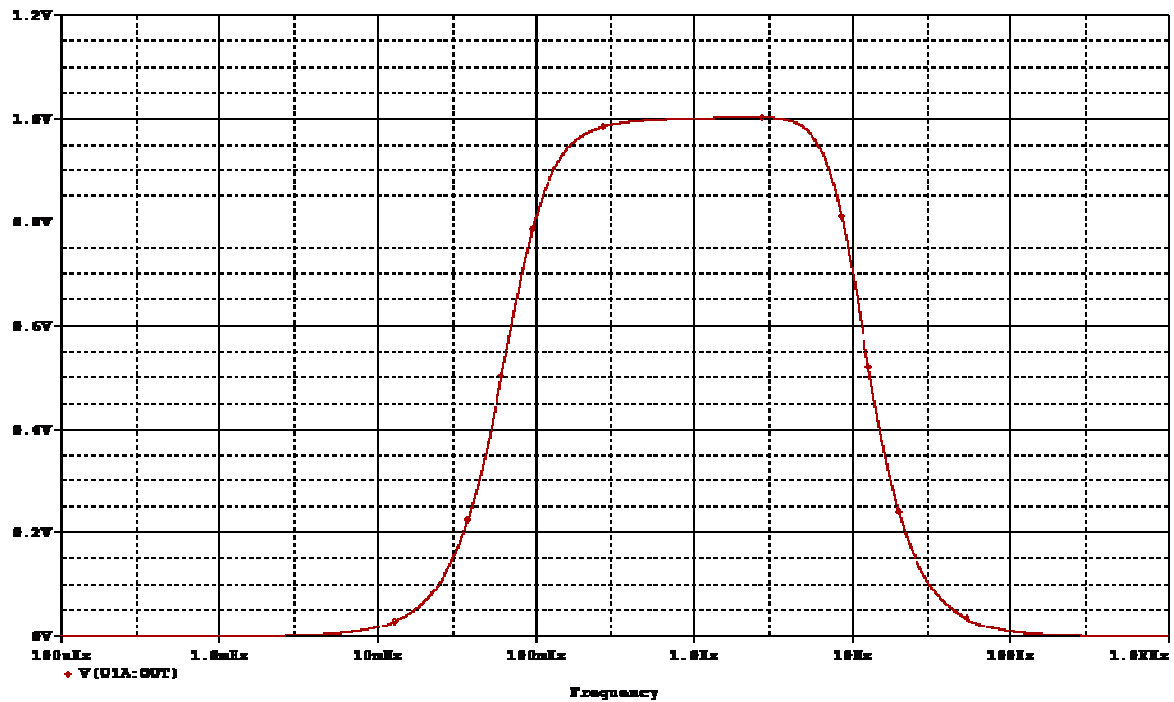


Figure 7. Output voltage of the band pass filter versus frequency.

Fig. 8 shows the amplified signals coming from the sensors for the red and infrared light respectively, after the continuous component removal.

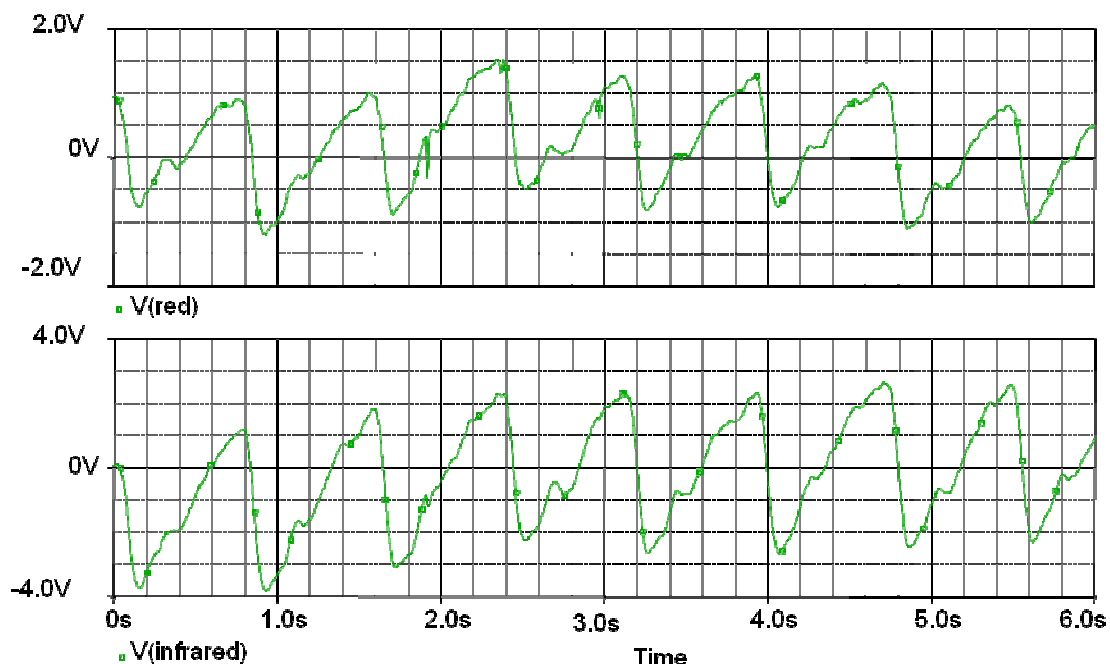


Figure 8. Amplified photoplethysmogram of signal after the continuous component removal.

The measurement of SpO_2 , as a function of the measured magnitude at the systolic and diastolic states on the two photoplethysmograms allows us to delete the dependence on the LED emissivity and on the photodiode sensibility. However, the relation (2) has to be empirically calibrated for the specific device [11].

Moreover, as our sensor has been shielded on three sides, we have been obtained a low probability that the ambient light may reach the photodiode and, therefore, may influence the measurement. Finally Fig. 9 shows the prototype realized and tested at the Electronic Devices Laboratory (Electrical and Electronic Department) of Polytechnic of Bari.

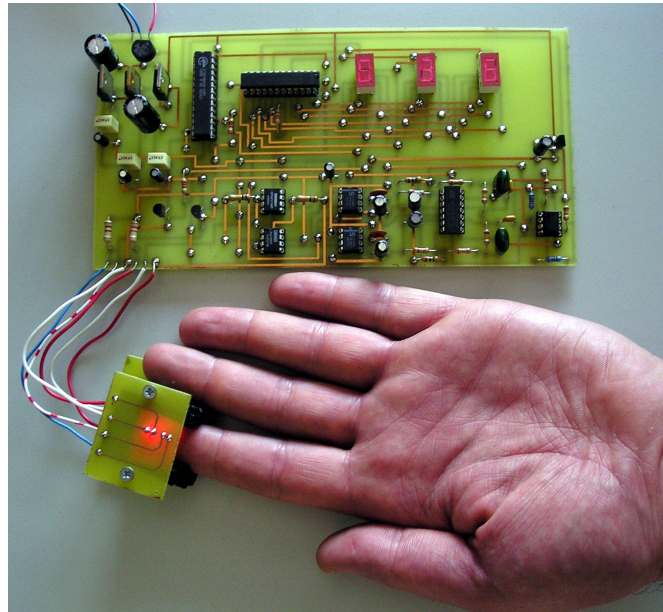


Figure 9. The realized prototype: a double-sided printed circuit board.

From this image it is clear that, even if the prototype is already small, but it will hugely squeezed using SMD (Surface Mount Device) technology. The image also shows the display with the measured values, there is no action to be taken to have the measurement since the device is continuously working.

IV. CONCLUSIONS AND FUTURE SCOPE

We have presented the design and realization of an electronic device for non invasive monitoring of the oxygen saturation in blood (transmission pulse oximeter). The main goals of our design have been the miniaturization, the cheapness and a good noise rejection. The key element to achieve these goals has been the PSoC, a system on chip family for mixed analogue and digital applications, programmable in both analogue and digital parts, allowing the implementation of a whole acquisition chain, from signal generator to the display driver, passing through sensor's amplifier, ADC and CPU. Having a single programmable device for both analogue and digital part, it has been easy to reach our goals. Furthermore this implementation of the pulse oximeter, using PSoC, has required some innovation in the circuit compared to previous schemes. The whole acquisition chain has a new plan that allows the collection of the continuous component of the signal. Moreover the whole data elaboration has been done in place using a local CPU, without requiring to pass data to an external computer.

For further development of this system, we are planning to include a miniaturized electrocardiograph.

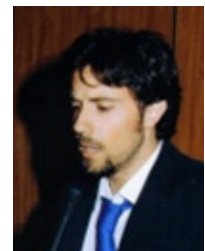
REFERENCES

- [1] Mengelkoc L. J., Martln D. & Lawler J., (1994) "A review of the Principles of Pulse Oximetry and Accuracy of Pulse Oximeter Estimates during Exercise", *Physical Therapy*, Vol. 74, No 1, pp. 40-49.
- [2] Datasheet PSoC CY8C27443, <http://www.cypress.com/?mpn=CY8C27443-24PXI>

- [3] Duun S., Haahr R. G., Birkelund K., Raahauge P., Petersen P., Dam H., Noergaard L. & Thomsen E. V., (2007) "A Novel Ring Shaped Photodiode for Reflectance Pulse Oximetry in Wireless Applications", *Proceedings of IEEE SENSORS Conference*, Atlanta, Georgia, USA, pp. 596-599.
- [4] Mannheimer P. D., Casciani J. R., Fein M. E. & Nierlich S. L., (1997) "Wavelength selection for low-saturation pulse oximetry", *IEEE Transactions on Biomedical Engineering*, Vol. 44, No 3, pp.148-158.
- [5] Konig V., Huch R. & Huch A., (1998) "Reflectance Pulse Oximetry – Principles and Obstetric Application in the Zurich System", *Journal of Clinical Monitoring and Computing*, Vol. 14, No 6, pp. 403-412.
- [6] Goldman J. M., Petterson M. T., Kopotic R. J. & Barker S. J., (2000) "Masimo Signal Extraction Pulse Oximetry", *Journal of Clinical Monitoring and Computing*, Vol. 16, No 7, pp. 475-483.
- [7] Gisiger P. A., Palma J. P. & Eberhard P., (2001) "Oxicarbo, a single sensor for the non-invasive measurement of arterial oxygen saturation and CO₂ partial pressure at the ear lobe", *Sensors and Actuators B: Chemical*, Vol. 76, No 1, pp.527 -530.
- [8] Mendelson Y. & Kent J. C., (1989) "Variations in optical absorption spectra of adult and fetal haemoglobins and its effect on pulse oximetry", *IEEE Transactions on Biomedical Engineering*, Vol. 36, No 8, pp. 844-849.
- [9] Fine I. & Weinreb A., (1995) "Multiple scattering effect in transmission pulse oximetry", *Medical & Biological Engineering & Computing*, Vol. 33, No 5, pp.709-712.
- [10] Sallen R. P. & Key E. L., (1955) "A Practical Method of Designing RC Active Filters", *IRE Transactions on Circuit Theory*, Vol. 2, No 1, pp.74–85.
- [11] Webster J. G., (1997) *Design of Pulse Oximeters*, IOP Publishing, Bristol, UK.

Authors

Roberto Marani received the Master of Science degree (cum laude) in Electronic Engineering in 2008 from Polytechnic of Bari where he is currently pursuing the Ph.D. degree. He worked in the Electronic Device Laboratory of Bari Polytechnic for the design, realization and testing of nanometrical electronic systems, quantum devices and FET on carbon nanotube. Moreover he worked in the field of design, modeling and experimental characterization of devices and systems for biomedical applications. Currently he is involved in the development of novel numerical models to study the physical effects involved in the interaction of electromagnetic waves with periodic metallic nanostructures. Dr. Marani has published over 50 scientific papers.



Gennaro Gelao received the Laurea degree in Physics from University of Bari, Italy, in 1993 and his Ph.D. degree in Physics in 1996 at CERN. He cooperates with the Electronic Device Laboratory of Polytechnic of Bari for the design, realization and testing of nanometrical electronic systems, quantum devices and CNTFETs. Dr. Gelao has published over 80 papers.



Anna Gina Perri received the Laurea degree cum laude in Electrical Engineering from the University of Bari in 1977. In the same year she joined the Electrical and Electronic Department, Polytechnic of Bari, where she is Professor of Electronics from 2002. Her current research activities are in the area of numerical modelling and performance simulation techniques of electronic devices for the design of GaAs Integrated Circuits and in the characterization and design of optoelectronic devices on PBG. Moreover she works in the design, realization and testing of nanometrical electronic systems, quantum devices and FET on carbon nanotube. Prof. Perri is the Director of Electron Devices Laboratory of the Electronic Engineering Faculty of Bari Polytechnic. She is author of over 250 book chapters, journal articles and conference papers and serves as referee for many international journals.



EFFECTS OF PGPR ON GROWTH AND NUTRIENTS UPTAKE OF TOMATO

Shahram Sharafzadeh

Department of Agriculture, Firoozabad Branch, Islamic Azad University, Firoozabad, Iran

ABSTRACT

Tomato is one of the most popular garden vegetable in the world. Tomatoes have high values in Vitamin A and C and are naturally low in calories. Inoculation with plant-growth promoting rhizobacteria (PGPR) has been attributed to the production of plant growth regulators at the root interface, which stimulate root development and result in better absorption of water and nutrients from the soil. A greenhouse experiment was conducted to evaluate the effects of some PGPR on growth and nutrients uptake of tomato (*Lycopersicon esculentum* Red Cherry) plants. Seven treatments were used for bacteria (*Pseudomonas*, *Azotobacter*, *Azosprillum*, *Pseudomonas* + *Azotobacter*, *Pseudomonas* + *Azosprillum*, *Azotobacter* + *Azosprillum* and *Pseudomonas* + *Azotobacter* + *Azosprillum*) which were compared to control. Plants were cut at prebloom stage. Maximum level of shoot fresh weight was shown on *Azotobacter* + *Azosprillum*, *Pseudomonas* + *Azotobacter* + *Azosprillum* and *Azosprillum* treatments which significantly differed from other treatments. Maximum level of root fresh weight was achieved in *Azotobacter* + *Azosprillum*, *Pseudomonas* + *Azotobacter* + *Azosprillum* and *Azotobacter* treatments which significantly differed from other treatments. Maximum level of shoot and root dry weights were achieved on *Azotobacter* + *Azosprillum* and *Pseudomonas* + *Azotobacter* + *Azosprillum* treatments. Minimum level of shoot and root dry weights were obtained in *Pseudomonas* + *Azosprillum*. Maximum root length was shown on *Azotobacter* + *Azosprillum* which significantly differed from other treatments. The highest amount of N, P and K were achieved on *Pseudomonas* + *Azotobacter* + *Azosprillum* treatment and the lowest amount was shown on *Pseudomonas* + *Azotobacter* treatment. Maximum level of Ca and Mg were obtained on *Pseudomonas* + *Azotobacter* and *Pseudomonas* + *Azosprillum* treatments which significantly differ from other treatments.

KEYWORDS: *Pseudomonas*, *Azotobacter*, *Azosprillum*, *Lycopersicum esculentum*

I. INTRODUCTION

Plant growth-promoting rhizobacteria (PGPR) help plants through different mechanisms, for example (i) the production of secondary metabolites such as antibiotics, cyanide, and hormonelike substances; (ii) the production of siderophores; (iii) antagonism to soilborne root pathogens; and (iv) phosphate solubilization [1,2,3,4,5,6,7]. These organisms possessing one or more of these characteristics are interesting since it may influence plant growth. Improvement of phosphorus (P) nutrition is one of the factors involved in plant growth promotion by PGPR. These bacteria may improve plant P acquisition by solubilizing organic and inorganic phosphate sources through phosphatase synthesis or by lowering the pH of the soil [8]. The objective of this study was to compare the effects of the PGPR at several treatments (alone and mixed) on growth and nutrients uptake of tomato plants.

II. MATERIALS AND METHODS

2.1. Plant Materials and Experimental Conditions

A greenhouse experiment was conducted to evaluate the effects of 7 treatments of bacteria (*Pseudomonas*, *Azotobacter*, *Azospirillum*, *Pseudomonas* + *Azotobacter*, *Pseudomonas* + *Azospirillum*, *Azotobacter* + *Azospirillum* and *Pseudomonas* + *Azotobacter* + *Azospirillum*) on tomato (*Lycopersicon esculentum* Red Cherry) growth and nutrients uptake. The plants were grown from seeds after inoculated with bacteria in pots containing 7 kg of field soil, sand and peat (1/3 ,v/v each of them). Experiment was set in a complete randomized design with four replicates. At prebloom stage, the shoots were cut at the soil surface level. The roots were separated from the soil. Shoot and root fresh weights and root length were measured, then dry weights of shoots and roots were determined after drying at 75°C.

2.2. Nutrient Determination

N, P and K were determined by kjeldahl, Olsen and flame photometry methods, respectively. Ca and Mg were determined by calciometry.

2.3. Statistical Analysis

Statistical analyses were done using SAS software. SAS (Statistical Analysis System) is an integrated system of software products provided by SAS Institute Inc. that enables programmers to perform statistical analysis. SAS is driven by SAS programs, which define a sequence of operations to be performed on data stored as tables. Means were compared by Duncan's multiple range test at $P < 0.05$ (5% level of probability).

III. RESULTS

The highest shoot fresh weight was observed in *Azotobacter* + *Azospirillum* (53.77 g/plant), *Pseudomonas* + *Azotobacter* + *Azospirillum* (53.29 g/plant) and *Azospirillum* (51.87 g/plant) treatments which significantly differed from other treatments. The lowest shoot fresh weight (42 g/plant) was obtained in *Pseudomonas* + *Azospirillum*. The maximum level of root fresh weight was achieved in *Azotobacter* + *Azospirillum* (10.81 g/plant), *Pseudomonas* + *Azotobacter* + *Azospirillum* (10.49 g/plant) and *Azotobacter* (10.30 g/plant) treatments which significantly differed from other treatments. Maximum level of shoot dry weight was shown on *Azotobacter* + *Azospirillum* (6.84 g/plant) and *Pseudomonas* + *Azotobacter* + *Azospirillum* (7.05 g/plant) treatments which significantly differed from others. The highest root dry weight was achieved on *Azotobacter* + *Azospirillum* (0.92 g/plant) and *Pseudomonas* + *Azotobacter* + *Azospirillum* (0.94 g/plant) treatments. Minimum level of shoot and root dry weights were achieved in *Pseudomonas* + *Azospirillum*. The maximum root length was shown on *Azotobacter* + *Azospirillum* (40.33 cm) which significantly differed from other treatments (Table 1).

Table 1. Effect of bacterial treatments on shoot and root fresh weights, shoot and root dry weights and root length.

Treatments	Shoot fw (g/plant)	Shoot dw (g/plant)	Root fw (g/plant)	Root dw (g/plant)	Root length (cm)
<i>Pseud.</i>	43.29b [†]	5.38cd	8.29c	0.63cd	34.18b
<i>Azoto.</i>	44.06b	5.46bcd	10.30a	0.79b	27.23c
<i>Azosp.</i>	51.87a	5.15d	8.59bc	0.58d	32.85b
<i>Pseud.</i> + <i>Azoto.</i>	43.75b	5.68bc	9.03b	0.60cd	32.13bc
<i>Pseud.</i> + <i>Azosp.</i>	42.00b	4.13e	7.59d	0.43e	31.40bc
<i>Azoto.</i> + <i>Azosp.</i>	53.77a	6.84a	10.81a	0.92a	40.33a
<i>Pseud.</i> + <i>Azoto.</i> + <i>Azosp.</i>	53.29a	7.05a	10.49a	0.94a	33.45b
Control	42.41b	5.93b	8.03cd	0.66c	34.00b

[†] In each column, means with the same letters are not significantly different at 5% level of Duncan's multiple range test.

The highest amount of N (32.65 mg/g dry matter), P (3.40 mg/g dry matter) and K (35.10 mg/g dry matter) were shown on *Pseudomonas* + *Azotobacter* + *Azospirillum* treatment which significantly differed from other treatments and the lowest amount was shown on *Pseudomonas* + *Azotobacter* treatment. The maximum level of Ca was achieved on *Pseudomonas* + *Azotobacter* (30.38 mg/g dry matter) and *Pseudomonas* + *Azospirillum* (30.30 mg/g dry matter) treatments which significantly differed from other treatments. The maximum level of Mg was observed on *Pseudomonas* + *Azotobacter* (6.18 mg/g dry matter) and *Pseudomonas* + *Azospirillum* (6.27 mg/g dry matter) treatments (Table 2).

Table 2. Effect of bacterial treatments on nutrients uptake in tomato.

Treatments	N (mg/g dry matter)	P (mg/g dry matter)	K (mg/g dry matter)	Ca (mg/g dry matter)	Mg (mg/g dry matter)
<i>Pseud.</i>	15.15 ^{ed}	2.28 ^{bc}	24.60 ^{bed}	20.05 ^c	4.30 ^c
<i>Azoto.</i>	16.70 ^d	2.23 ^{bc}	26.6 ^{bc}	24.93 ^b	5.15 ^b
<i>Azosp.</i>	24.45 ^b	2.55 ^b	28.70 ^b	22.05 ^{bc}	5.18 ^b
<i>Pseud</i> + <i>Azoto.</i>	10.93 ^e	1.93 ^c	21.23 ^d	30.38 ^a	6.18 ^a
<i>Pseud</i> + <i>Azosp.</i>	18.53 ^{cd}	2.08 ^{bc}	20.73 ^d	30.30 ^a	6.27 ^a
<i>Azoto.</i> + <i>Azosp.</i>	24.25 ^b	2.30 ^{bc}	24.10 ^{bed}	21.90 ^{bc}	5.45 ^b
<i>Pseud</i> + <i>Azoto.</i> + <i>Azosp.</i>	32.65 ^a	3.40 ^a	35.10 ^a	21.40 ^{bc}	4.05 ^c
Control	22.15 ^{bc}	2.35 ^{bc}	22.60 ^{cd}	22.20 ^{bc}	4.43 ^c

[†] In each column, means with the same letters are not significantly different at 5% level of Duncan's multiple range test.

IV. DISCUSSION

The results indicated that PGPR affect the growth and nutrients uptake. In the impact of root inoculation with beneficial rhizosphere microorganisms on some quality parameters is being explored [9,10,11].

Facilitating plant nutrition could be the mechanism by which PGPR enhance crop yield, since the nutritional plants status is enhanced by increasing the availability of nutrients in the rhizosphere [12,13].

Phytohormones produced by PGPR, are believed to be changing assimilate partitioning patterns in plants altering growth in roots, the fructification process and development of the fruit under production conditions [14].

This work supports that tomato root inoculation with PGPR enhances growth under greenhouse conditions. However, field experiments should be carried out to ensure that positive effects are maintained under conventional production systems.

A series of other factors (ability to grow on root exudates, to synthesize amino acids and vitamins) defined as "rhizospheric competence" is involved in the establishment of effective and enduring root colonization by an introduced bacterium [15].

Pseudomonas fluorescens 92rk, alone or co-inoculated with P190r, increased mycorrhizal colonization of tomato roots by *G. mosseae* BEG12. This result suggests that strain 92rk behaves as a mycorrhiza helper bacterium (MHB) in *L. esculentum*. MHB have been described for ectomycorrhizal symbiosis [16] and only a few examples of MHB have been reported for AM symbiosis [17,18]. *P. fluorescens* 92rk increased total root length, surface area and volume. This is in agreement with the effects of *P. fluorescens* A6RI [19] and 92r [20] on the development of tomato and cucumber root, respectively. Longer root systems are more adapted to soil exploration and exploitation [21]. The

modification of root architecture parameters induced by the bacterial strains could be related to increased P acquisition: root systems with higher root surface area and volume are indeed characterized by a higher absorptive surface.

An investigation showed the effects of inoculating of two Bred cultivars of tomato (F1 Hybrid, Delba and F1 Hybrid, Tivi) roots with plant growth-promoting rhizobacteria (PGPR). *Azotobacter* was more effective than *Pseudomonas* to increase all traits' value except for shoot dry weight and K Content [22]. Differences between genotypes can explain differences between results.

An investigation showed that PGPR and AMF (fungus) can increase tomato fruit quality. It may be related to increasing of minerals by inoculated plants [23].

Increased nutrient uptake by plants inoculated with plant growth promoting bacterium has been attributed to the production of plant growth regulators at the root interface, which stimulated root development and resulted in better absorption of water and nutrients from the soil [24,25,26].

V. CONCLUSION

In conclusion, *Azotobacter* + *Azospirillum* and *Pseudomonas* + *Azotobacter* + *Azospirillum* resulted in the highest values of shoot fresh and dry weights and root fresh and dry weights at prebloom stage. *Pseudomonas* + *Azotobacter* + *Azospirillum* treatment was the best for N, P and K uptake in tomato shoots.

REFERENCES

- [1] A.W. Bakker and B. Schippers, "Microbial cyanide production in the rhizosphere in relation to potato yield reduction and *Pseudomonas spp.*-mediated plant growth-stimulation", Soil Biol. Biochem., Vol. 19, PP. 451-457, 1987.
- [2] P.J. Dart, "Nitrogen fixation associated with non-legumes in agriculture", Plant Soil, Vol. 90, PP. 303-334, 1986.
- [3] A.N. Dubeikovsky, E.A. Mordukhova, V.V. Kochetkov, F.Y. Polikarpova, and A.M. Boronin, "Growth promotion of blackcurrant softwood cuttings by recombinant strain *Pseudomonas fluorescens* BSP53a synthesizing an increased amount of indole-3-acetic acid", Soil Biol. Biochem., Vol. 25, PP. 1277-1281, 1993.
- [4] A.H. Goldstein, "Bacterial solubilization of mineral phosphates: historical perspective and future prospects", Am. J. Altern. Agric., Vol. 1, PP. 51-57, 1986.
- [5] J. Leong, "Siderophores: their biochemistry and possible role in the biocontrol plant pathogens", Annu. Rev. Phytopathol., Vol. 24, PP. 187-209, 1986.
- [6] M.N. Schroth and J.G. Hancock, "Selected topics in biological control". Annu. Rev. Microbiol., Vol. 35, PP. 453-476, 1981.
- [7] D.M. Weller, "Biological control of soilborne pathogens in the rhizosphere with bacteria", Annu. Rev. Phytopathol., Vol. 26, PP. 379-407, 1988.
- [8] H. Rodriguez and R. Fraga, "Phosphate solubilizing bacteria and their role in plant growth promotion", Biotechnol. Adv., Vol. 17, PP. 319-339, 1999.
- [9] G. Charron, V. Furlan, M. Bernier-Cardou and G. Doyon, "Response of onion plants to arbuscular mycorrhizae. 1. Effects of inoculation method and phosphorus fertilization on biomass and bulb firmness", Mycorrhiza, Vol. 11, PP. 187-197, 2001.
- [10] C. Kaya, D. Higgs, H. Kirnak and I. Tas, "Mycorrhizal colonization improves fruit yield and water use efficiency in watermelon (*Citullus lanatus* Thunb.) grown under well-watered and water stressed conditions", Plant Soil, Vol. 253, PP. 287-292, 2003.
- [11] H.G. Mena-Violante, O. Ocampo-Jimenez, L. Dendooven, G. Martinez-Soto, J. Gonzalez-Castafieda, F.T. Davies Jr. and V. Olalde-Portugal, "Arbuscular mycorrhizal fungi enhance fruit growth and quality of chile ancho (*Capsicum annuum* L. cv San Luis) plants exposed to drought", Mycorrhiza, Vol. 16, PP. 261-267, 2006.
- [12] E. Bar-Ness, Y. Hadar, Y. Chen, V. Romheld, and H. Marschner, "Short term effects of rhizosphere microorganisms on Fe uptake from microbial siderophores by maize and oat", Plant Physiol., Vol. 100, PP. 451-456, 1992.
- [13] A.E. Richardson, "Prospects for using soil microorganisms to improve the acquisition of phosphorus by plants", Aust. J. Plant Physiol., Vol. 28, PP. 897-906, 2001.
- [14] J.A. Lucas-Garcia, A. Probanza, B. Ramos, M. Ruiz-Palomino and F.J. Gutierrez Manero, "Effect of inoculation of *Bacillus licheniformis* on tomato and pepper", Agronomie, Vol. 24, PP. 169-176, 2004.

- [15] B.J.J. Lugtenberg and L.C. Dekkers, "What makes *Pseudomonas* bacteria rhizosphere competent?" *Environ. Microbiol.*, Vol. 1, PP. 9-13, 1999.
- [16] J. Garbaye, "Helper bacteria: a new dimension to the mycorrhizal symbiosis", *New Phytol.*, Vol. 128, PP. 197-210, 1994.
- [17] M. Toro, R. Azcn and J.M. Barea, "Improvement of arbuscular development by inoculation of soil with phosphate-solubilizing rhizobacteria to improve rock phosphate bioavailability (32P) and nutrient cycling", *Appl. Environ. Microbiol.*, Vol. 63, PP. 4408-4412, 1997.
- [18] S. Singh and K.K. Kapoor, "Effects of inoculation of phosphate solubilizing microorganisms and an arbuscular mycorrhizal fungus on mungbean grown under natural soil conditions", *Mycorrhiza*, Vol. 7, PP. 249-253, 1998.
- [19] E. Gamalero, M.G. Martinotti, A. Trotta, P. Lemanceau and G. Berta, "Morphogenetic modifications induced by *Pseudomonas fluorescens* A6RI and *Glomus mosseae* BEG12 in the root system of tomato differ according to plant growth conditions", *New Phytol.*, Vol. 155, PP. 293-300, 2002.
- [20] E. Gamalero, L. Fracchia, M. Cavaletto, J. Garbaye, P. Frey-Klett, G.C. Varese and M.G. Martinotti, "Characterization of functional traits of two fluorescent pseudomonads isolated from basidiomes of ectomycorrhizal fungi", *Soil Biol. Biochem.*, Vol. 35, PP. 55-65, 2003.
- [21] G. Berta, A. Fusconi and J.E. Hooker, "Arbuscular mycorrhizal modifications to plant root systems", In: S. Gianinazzi and H. Schuepp (eds) "Mycorrhizal Technology: from genes to bioproducts achievement and hurdles in arbuscular mycorrhizal research", Birkh_user, Basel, pp. 71-101, 2002.
- [22] M. Zare, K. Ordoorkhani and O. Alizadeh, "Effects of PGPR and AMF on Growth of Two Bred Cultivars of Tomato", *Adv. Environ. Biol.*, Vol. 5, PP. 2177-2181, 2011.
- [23] K. Ordoorkhani, K. Khavazi, A. Moezzi and F. Rejali, "Influence of PGPR and AMF on antioxidant activity, lycopene and potassium contents in tomato", *Afr. J. Agric. Res.*, Vol. 5, PP. 1108-1116, 2010.
- [24] J.W. Kloepper, R.M. Zablowicz, B. Tipping and R. Lifshitz, "Plant growth mediated by bacterial rhizosphere colonizers", In: D.L. Keister and B. Gregan (eds.) "The rhizosphere and plant growth", 14. BARC Symposium, PP. 315-326, 1991.
- [25] W. Zimmer, K. Kloos, B. Hundeshagen, E. Neiderau and H. Bothe, "Auxin biosynthesis and enitrification in plant growth promotion bacteria", In: J. Fendrik, J. De Gallo Vandeleyden and D. De Zamoroczy (eds.) "Azospirillum VI and related microorganisms", Series G:Ecological, Vol. 37, PP. 120 141, 1995.
- [26] G. Hoflich and G. Kuhn, "Forderung das Wachstums und der Nahrstoffaufnahme bei kurziferen O- und Zwischenfruchten durch inokulierte Rhizospherenmikroorganismen", *Zeischrift fu r Pflanzenerna hrung und Bodenkunde*, Vol. 159, PP. 575-578, 1996.

Author

Shahram Sharafzadeh was born in Shiraz, Iran in 1971. He received his Bachelor degree in Horticultural Science from Shiraz University, Iran in 1994; MSc in Horticultural Science from Shiraz University, Iran in 1998 and Ph.D in Horticultural Science from Science and Research Branch, Islamic Azad University, Tehran, Iran in 2008. He is working as a full time Lecturer, assistant professor, in the Firoozabad Branch, Islamic Azad University, Firoozabad, Iran. His research interests include medicinal and aromatic plants and biotechnology. He is supervisor and advisor of some MSc thesis. There are several projects he is working on such as effects of organic matters on growth and secondary metabolites of plants.



THE APPLICATION OF PSO TO HYBRID ACTIVE POWER FILTER DESIGN FOR 3 PHASE 4-WIRE SYSTEM WITH BALANCED & UNBALANCED LOADS

B. Suresh Kumar¹, K. Ramesh Reddy² & S. Archana³

^{1&3}Department of Electrical and Electronics Engineering, CBIT, Hyderabad, India

²Department of Electrical and Electronics Engineering, GNITS, Hyderabad, India

ABSTRACT

This paper presents a application of PSO to Hybrid active power filter used to compensate for total harmonic distortion in three-phase four-wire systems. The shunt active filter employs a simple method for the calculation of the reference compensation current based on Fast Fourier Transform. The presented Shunt Active Power filter is able to operate in balanced, unbalanced and Variable load conditions. Classic filters may not have satisfactory performance in fast varying conditions. But auto tuned active power filter gives better results for harmonic minimization, reactive power compensation and power factor improvement. The proposed auto tuned shunt active filter maintains the THD well within IEEE-519 standards. The proposed methodology is extensively tested for wide range of different Loads with Improved dynamic behavior of shunt active power filter using PSO to Hybrid active power filter. The results are found to be quite satisfactory to mitigate harmonic Distortions, reactive power compensation and power factor correction thereby increase in Power Quality improvement and reduction in %THD.

KEYWORDS: Hybrid active power filter (HAPF), Multiobjective optimization, particle swarm optimization (PSO), Total Harmonic distortion (THD), Power factor, Reactive power

I. INTRODUCTION

Power Systems have to cope with a variety of nonlinear Loads which introduce significant amounts of harmonics. IEEE Standard 519-1992 provides a guideline for the limitation and mitigation of harmonics. Passive power filters (PPFs), Active power filters (APFs), and hybrid active power filters (HAPFs) can all be used to eliminate harmonics. For Medium- and high-voltage systems, PPFs and HAPFs appear to be better choices considering cost where the ratings are of several tens of megavolt-ampere. The design of such PPFs and HAPFs is a complicated nonlinear programming problem. Conventional trial-and-error Methods based on engineering experience are commonly used, but the results are not optimal in most cases.

In recent years, many studies have appeared involving optimal PPF design. A Method based on the sequential unconstrained minimization Technique has been used for PPF design because of its simplicity and versatility, but numerical instability can limit the application of this method. PPF design using simulated Annealing has been reported, but the major drawback is the repeated annealing.

Genetic algorithms (gas) have been widely used in PPF design, but the computing burden and convergence problems are disadvantages of this approach. A design method for PPFs using a hybrid

Differential evolution Algorithm has also been proposed, but the algorithm is Complex, involving mutation, crossover, migrant, and accelerated

Operations For the optimal design of HAPFs, a method based on gas has been proposed in order to minimize the rating of APF, but no other optimal design methods appear to have been suggested. Many methods treated the optimal design of PPFs and HAPFs as a single objective problem. In fact, filter Design should determine the optimal solution where there are multiple objectives. As these objectives generally conflict with One another, they must be cautiously coordinated to derive a Good compromise solution.

In this paper, optimal multi objective designs for both PPFs and HAPFs using an advanced particle swarm optimization (PSO) algorithm are reported. The objectives and constraints were developed from the viewpoint of practicality and the Filtering characteristics.

For the optimal design of PPFs, the capacity of reactive Power compensation, the original investment cost, and the total Harmonic distortion (THD) were taken as the three objectives. The constraints included individual harmonic distortion, fundamental Reactive power compensation, THD, and parallel and Series resonance with the system. For the optimal design of HAPFs, the capacity of the APF, The reactive power compensation, and the THD were taken as the three objectives; the constraints were as for the PPFs.

The Uncertainties of the filter and system parameters, which will Cause detuning, were also considered as constraints during the optimal design process. A PSO-based algorithm was developed to search for the optimal solution. The numerical results of case Studies comparing the PSO method and the conventional trial and- Error method are reported. From which, the superiority and Availability of the PSO method and the designed filters are certified.

II. SYSTEM UNDER STUDY

A typical 10-kV 50-Hz system with nonlinear loads, as shown in Fig. 1, was studied to determine the optimal design for both PPFs and HAPFs. The nonlinear loads are the medium frequency furnaces commonly found in steel plants with abundant harmonic currents, particularly the fifth and seventh orders, as shown in Table I. The utility harmonic tolerances given in IEEE Standard 519-1992 and the Chinese National Standard GB/T14549-93 are listed in Table I as percentages of the fundamental current.

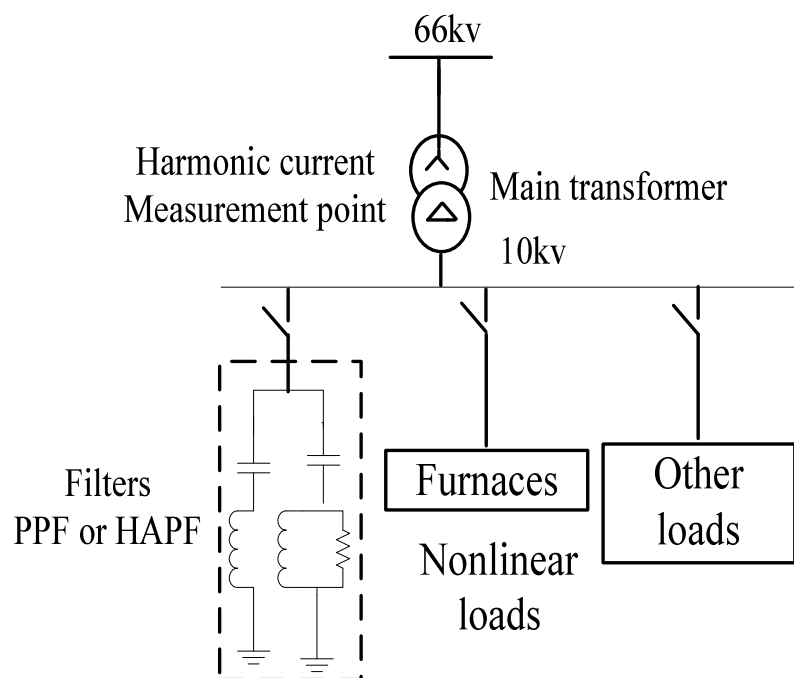


Fig 1. Single diagram of system for case studies.

Table I : Harmonic Current distributions in phase A and utility tolerances

Harmonic Order	Measured Value (%)	National Standard (%)	IEEE standard 519-1992 (%)
5	6.14	2.61	4
7	2.77	1.96	4
11	1.54	1.21	2
13	0.8	1.03	2
17	0.6	0.78	1.5
19	0.46	0.7	1.5
23	0.95	0.59	0.6
25	0.93	0.53	0.6
THD	7.12	5	5

Table I shows that current THD, and the 5th, 23rd, and 25th order harmonic currents exceed the tolerances based on both standards. In addition, the 7th and 11th order harmonics exceed the tolerance based on the National standard.

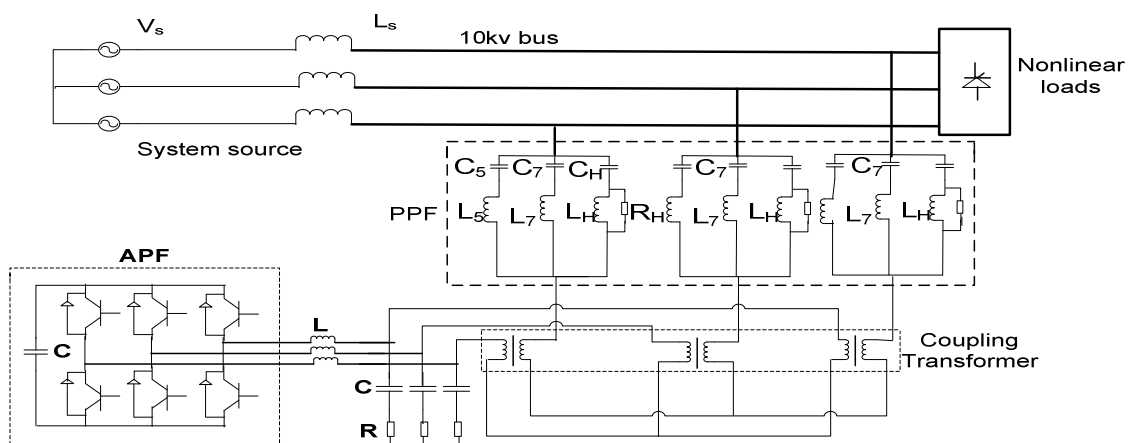
Filters must therefore be installed to mitigate the harmonics sufficiently to satisfy both standards. Both PPF and HAPF are suitable and economical for harmonic mitigation in such systems. For this system with nonlinear loads as medium frequency furnaces, the even and triple harmonics are very small and far below the standard values, so these harmonics are not considered. In addition, the harmonic voltages are in fact very small, so the voltages are assumed to be ideal.

The fundamental current and reactive power demands are 1012 A and 3–4 MVar, respectively. The short circuit capacity is 132 MVA, and the equivalent source inductance of the system is 2.4 mH

III. HAPF DESIGN BASED ON PSO

A. HAPF Structure and Performance:

In order to demonstrate the optimal design method of HAPFs based on PSO, an HAPF was designed and is shown in Fig. 2; it is supposed to be used in the same situation as that shown in Fig. 1. In this HAPF, PPFs are mainly used to compensate for harmonics and reactive power, and an APF is used to improve the filtering performance. The PPF consists of the fifth and seventh single-tuned filters and a high-pass damped filter. The APF is implemented with a three-phase voltage-source inverter. Fig. 3(a) shows the single-phase equivalent circuits of the HAPF, assuming that the APF is an ideal controllable voltage VAF and that the load is an ideal current source IL . ZS is the source impedance, ZF is the total impedance of the PPF, V_{pcc} is the voltage of the injected point, and K is the controlling gain of the APF.

**Fig.2.** Shunt HAPF.

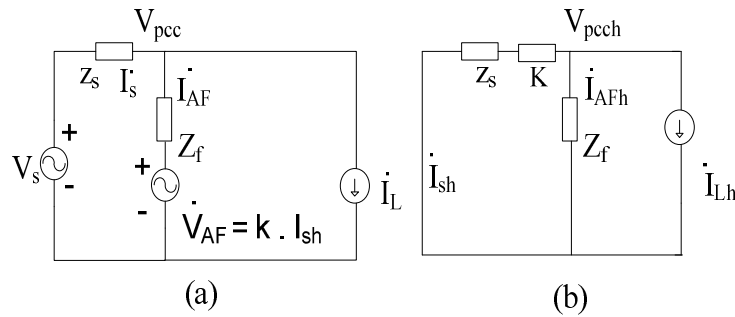


Fig 3.Single-phase equivalent circuits of the HAPF (a) Equivalent circuit.
(b) Equivalent harmonic circuit.

The equivalent harmonic circuit is redrawn as in Fig. 3(b). The harmonic current I_{sh} into the system source and the harmonic attenuation factor γ are given in the following equations:

$$I_{sh} = \left| \frac{Z_F}{K + Z_F + Z_S} \right| I_{Lh} \rightarrow (1)$$

$$\gamma = \frac{I_{sh}}{I_{Lh}} = \left| \frac{Z_F}{K + Z_F + Z_S} \right| \rightarrow (2)$$

Assuming that the fundamental component of the supply voltage is fully dropped across the PPF, the voltage and current of APF can be derived as follows [24]:

$$V'_{AF} = \sum_h V'_{AFh} = \sum_h -Z_{Fh} \cdot I'_{AFh} = \sum_h Z_{Fh} \cdot I'_{Lh} \rightarrow (3)$$

$$I'_{AF} = I'_{AF1} + \sum_h I'_{AFh} \rightarrow (4)$$

The rms value of V_{AF} is defined as

$$V_{AF} = \sqrt{\sum_{h=5,7,11,13,17,19,23,25} V_{AFh}^2} \rightarrow (5)$$

The capacity of the APF is determined by the current I'_{AF} and the voltage V_{AF} . It is obvious that the low VA rating of APF can be achieved by decreasing I'_{AF} and V_{AF} . In this shunt hybrid topology, the current I'_{AF} is almost constant, so the only way to reduce the VA rating of APF is to decrease the voltage V_{AF} .

B. Multi objective Optimal Design Model for HAPF:

As mentioned earlier, when designing an HAPF, it is very important to minimize the capacity of the APF component, and there are some other objectives and constraints to be considered when the APF of the HAPF is cut off due to faults, the PPF part keeps work in to mitigate harmonics until the APF is restored. It follows that some additional constraints should be included in respect of such occasions. The constructions of objective functions and constraints are described next.

Three important objective functions are defined as follows.

1) Minimize the capacity of the APF, which is mainly determined by the harmonic voltage across it

$$\min V_{AF} = \sqrt{\sum_{h=5,7,11,13,17,19,23,25} V_{AFh}^2} \rightarrow (6)$$

2) Minimize the current THD with HAPF

$$\min THDI_{HAPF} = \sqrt{\sum_{h=2}^N \left(\frac{I_{sh}}{I_1} \right)^2} \rightarrow (7)$$

where $THDI_{HAPF}$ is the current THD with the HAPF in place; and the definitions of I_{sh} , I_1 , and N are the same as those in (7).

3) Maximize the reactive power compensation

$$\max \sum_{i=5,7,H} Q_i \rightarrow (8)$$

Where Q_i is same with that in (9)

Constraints are defined as follows.

The tolerated values quoted hereinafter are also based on the National standard.

1) Requirements of total harmonic filtering with the HAPF

$$THDI_{HAPF} \leq THDI_{MAX} \rightarrow (9)$$

Where $THDI_{MAX}$ is defined in (10)

When PPF runs individually with the APF cutoff, an additional constraint is added as follows:

$$THDI_{PPF} \leq THDI_{MAX} \rightarrow (10)$$

Where $THDI_{PPF}$ is the THD with the PPF alone

2) Requirements of individual harmonic filtering with HAPF and PPF alone: Each order harmonic should satisfy the standards, so the following constraints are included:

$$I_{HAPF_{sh}} \leq I_{h_{max}}, h=5,7,11,13,17,19,23,25 \rightarrow (11)$$

$$I_{PPF_{sh}} \leq I_{h_{max}}, h=5,7,11,13,17,19,23,25 \rightarrow (12)$$

Where $I_{HAPF_{sh}}$ and $I_{PPF_{sh}}$ are, respectively, the rms values of the h_{th} order harmonic current into the system source with the HAPF and the PPF alone $I_{h_{max}}$ is defined by (11).

3) Fundamental reactive power compensation limits: The fundamental reactive power must be restricted as

$$Q_{min} \leq \sum_{i=5,7,H} Q_i \leq Q_{max} \rightarrow (13)$$

Where Q_{min} and Q_{max} are as defined in (12).

4) Parallel and series resonance restrictions: Parallel and series resonance with system source will rarely happen for the HAPF due to the active damping function of the APF. Nevertheless, it is necessary to consider, during the HAPF design, parallel and series resonance restrictions when PPF works alone with the APF cutoff. Therefore, constraints are constructed, which are the same as those constructed during the PPF optimal design in (13)–(16).

5) Consideration of the detuning constraints: The HAPF system is not sensitive to detuning effects because of the APF damping function. In the worst case, that the system impedance decreases by 20%, the system frequency changes to 50.5 Hz, the capacitance of each branch increases by 5%, and the reactance also increases by 2%, then the filtering performance of the PPF alone should still satisfy all the standards and limit described earlier, as set out in (10), (12), and (13).

C. Optimal Design for HAPF Based on PSO Based on the objectives and constraints constructed earlier for HAPF, the multi objective optimization task is carried out using an advanced PSO algorithm. The capacitance in each branch of the PPF and the characteristic frequency of the high-pass damped filter are chosen as optimal variables

$X_i = (C5, C7, CH, fH)T$, while the tuning frequencies of the fifth and seventh single-tuned filters are predetermined as 242 and 342 Hz, respectively. According to the optimization objectives, the corresponding fitness functions are defined as

$$F_1'(X) = V_{AF} \rightarrow (14)$$

$$F_2'(X) = THDI_{HAPF} \rightarrow (15)$$

$$F_3'(X) = \sum_{i=5,7,H} Q_i \rightarrow (16)$$

Similar methods were adopted to solve this multi objective optimization problem. The objective of minimizing the APF capacity is chosen as the final objective, while the other two objectives are solved by using acceptable level constraints, as shown in the following equations:

$$\min F_1' \rightarrow (17)$$

$$F_2' \leq \alpha_1' \rightarrow (18)$$

$$\alpha_2' \leq F_3' \leq \alpha_3' \rightarrow (19)$$

Where α_1' , α_3' , and α_2' are the highest and lowest acceptable levels for the secondary objectives, respectively. The overall optimization process for the HAPF based on PSO is similar to that of the PPF in Fig. 4.

Table II : Design results of HAPFs based on PSO and conventional methods

Design parameters	Pso-method	Conventional method
The 5th Single-tuned filter	$C_5=59.76\mu\text{F}$ $L_5=7.24\text{mH}$ $Q_5=60$	$C_5=80.6\mu\text{F}$ $L_5=5.37\text{mH}$ $Q_5=60$
The 7th single -tuned filter	$C_7=12.32\mu\text{F}$ $L_7=17.58\text{mH}$ $Q_7=60$	$C_7=23.76\mu\text{F}$ $L_7=9.11\text{mH}$ $Q_7=60$
High-pass damped filter	$C_H=52.06\mu\text{F}$ $L_H=1.20\text{mH}$ $m=0.5$	$C_H=15.28\mu\text{F}$ $L_H=3.32\text{mH}$ $m=0.5$

Table III : Harmonic current distributions with HAPFs based on PSO and conventional methods

Harmonic orders	PSO Method (%)	Conventional Method (%)
5	0.24	0.17
7	0.24	0.11
11	0.25	0.71
13	0.1	0.3
17	0.07	0.16
19	0.06	0.12
23	0.13	0.26
25	0.13	0.26
THD	0.48	0.91
V_{AF}	116.64 V	340.82 V
Reactive Power Compensation	4MVar	3.88MVar

The design results of HAPFs using PSO and conventional trial-and-error methods are listed in Table II. The design results based on the conventional method in Table II. It can be seen that the harmonic currents and reactive power are well compensated by both HAPFs and that the HAPF designed using the method based on PSO can obtain better filtering performance with lower THD (0.48%) and larger reactive power compensation (4MVar). Moreover, the voltage VAF of the APF, in this case, was much smaller than that based on conventional method. Therefore, the investment cost of the whole system is much reduced. Table IV shows the harmonic current distributions when the PPF is working alone, without the APF.

A comparison is made between the PSO method and conventional method, and it can be seen that all the harmonic currents and the THD are still within the standards, and the filtering performance of PPF based on PSO is a little better.

Table IV: Harmonic current distributions with PPFs alone Based on PSO and conventional methods

Harmonic orders	PSO Method (%)	Conventional Method (%)
5	1.1	0.82
7	0.76	0.39
11	0.94	1.13
13	0.26	0.60
17	0.14	0.29
19	0.11	0.21
23	0.21	0.40
25	0.20	0.38
THD	1.68	1.71

Table V: Harmonic current distributions with HAPFs alone considering detuning effects

Harmonic Orders	PSO Method HAPF (%)	Conventional Method HAPF (%)
5	0.65	0.44
7	0.75	0.27
11	0.23	0.71
13	0.1	0.27
17	0.08	0.16
19	0.06	0.13
23	0.14	0.28
25	0.14	0.28
THD	1.05	1.02

In order to verify the filtering performances of HAPF alone under the worst detuning situations, comparisons are shown in Table V. It is clear that both HAPFs, using PSO method and conventional method, can obtain excellent filtering performance in spite of detuning effects.

Fig. 4 shows the harmonic attenuation factors of HAPF alone using the PSO design method and considering detuning effects. It can be seen that the harmonic currents are still well attenuated, and no resonance point can be found. Furthermore, the attenuation factor of HAPF is much smaller than that of PPF, which shows the excellent harmonic mitigation performance of HAPF.

The simulation using the MATLAB/SIMULINK software has been run based on field measurement data. The compensated source current with the HAPF is shown in Fig. 5. From Fig. 5, we can see that the source current is very close to a pure sine wave, with the current THD decreasing to 0.48%.

Fig. 6 shows the convergence characteristics of the PSO algorithm developed in this paper for optimal design of HAPF. In this paper, the PSO algorithm is run 200 times, and every time, it can converge within 360 iterations. All those can demonstrate the efficiency and validity of PSO for the optimal HAPF design.

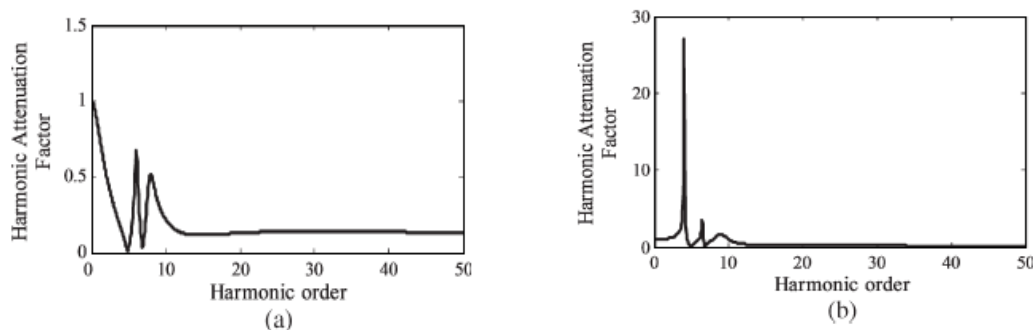


Fig 4. Harmonic attenuation factors of the HAPF and its PPF alone based on the PSO method.

- (a) Harmonic attenuation factor of the HAPF based on the PSO method.
(b) Harmonic attenuation of the PPF alone based on the PSO method.

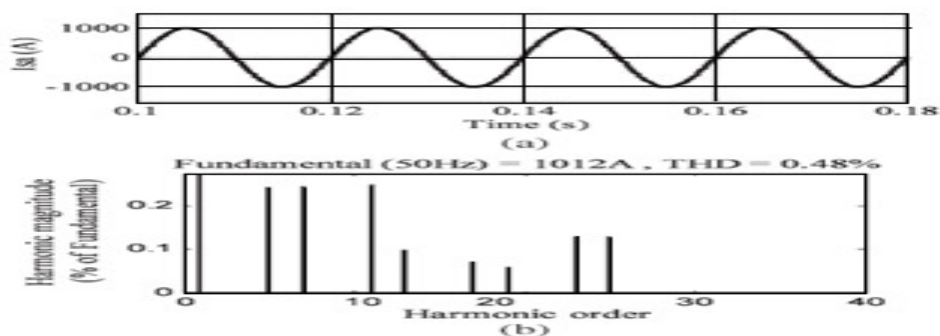


Fig 5.Compensated source current and its THD analysis with HAPF based on the PSO method

- (a) Compensated source currents of phase A with HAPF.
(b) THD analysis of compensated source current.

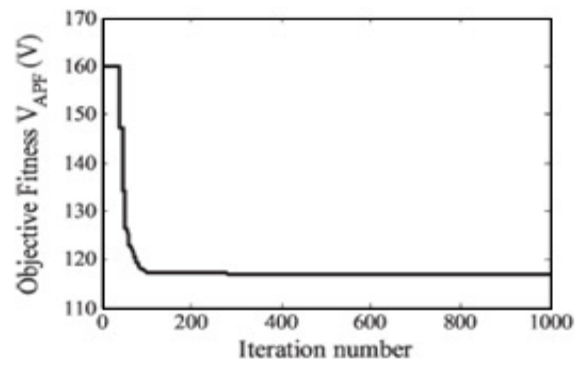


Fig.6. Convergence characteristics of PSO for HAPF design.

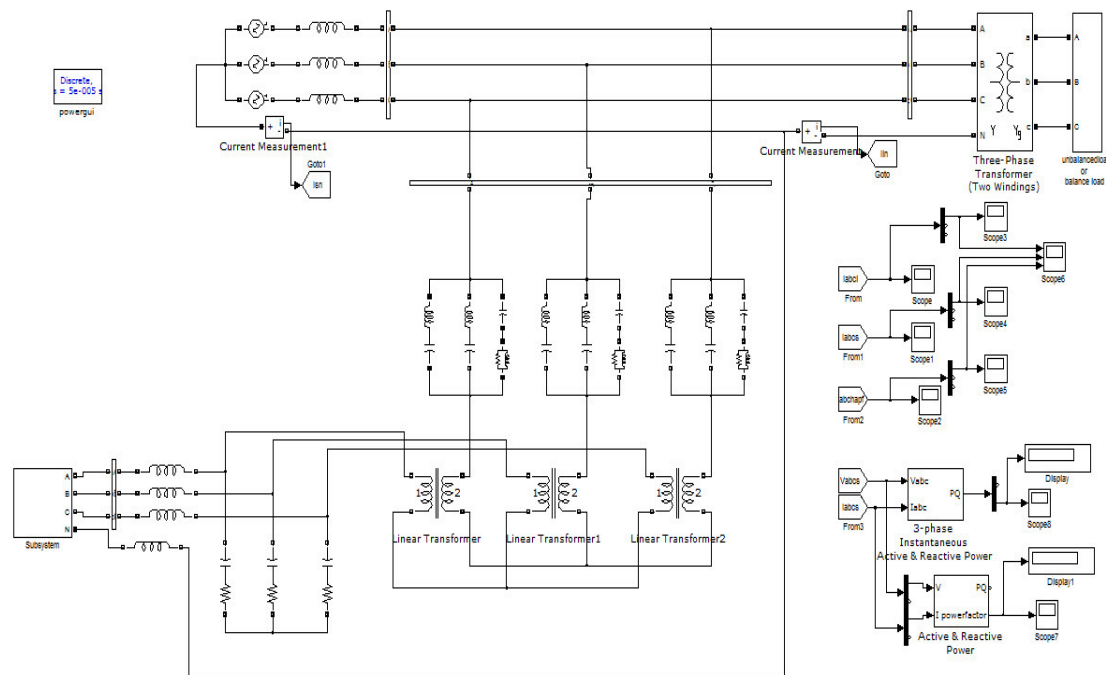


Fig.7. Simulink model of HAPF using PSO & without PSO with balanced, & unbalanced models

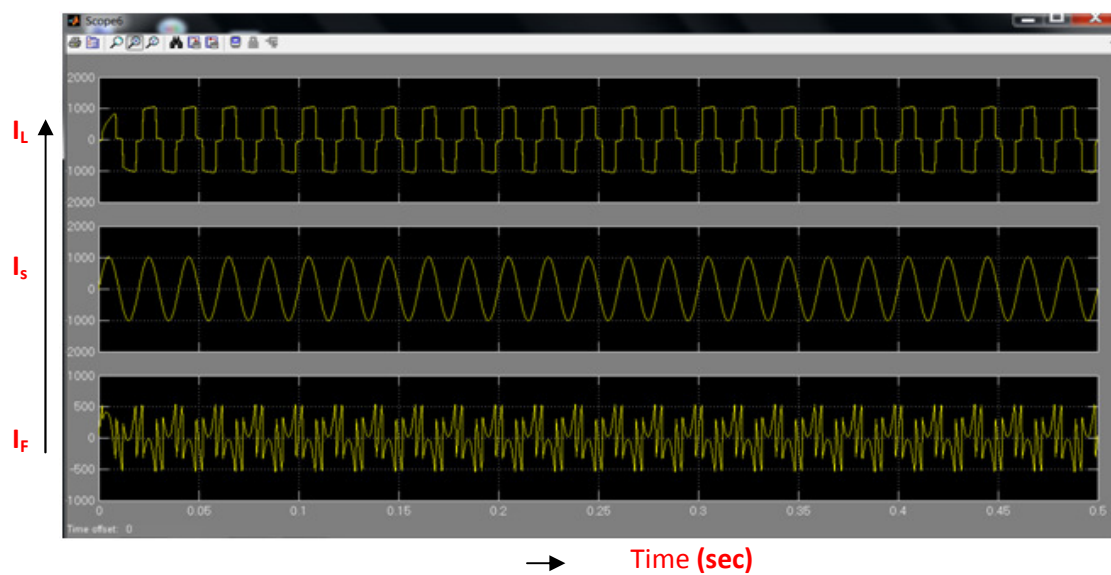


Fig 8. Wave form of balanced load for HAPF-Conventional Method

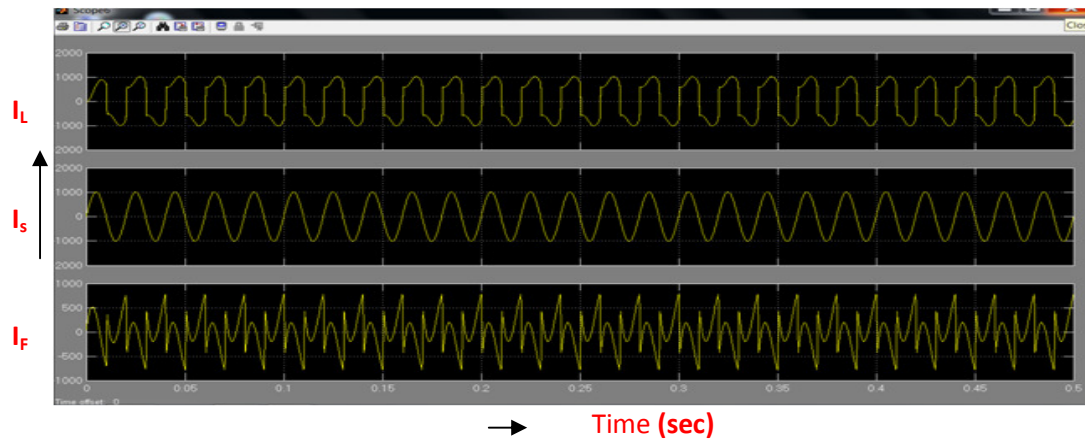


Fig.9. Wave form of Unbalanced load for HAPF-Conventional Method

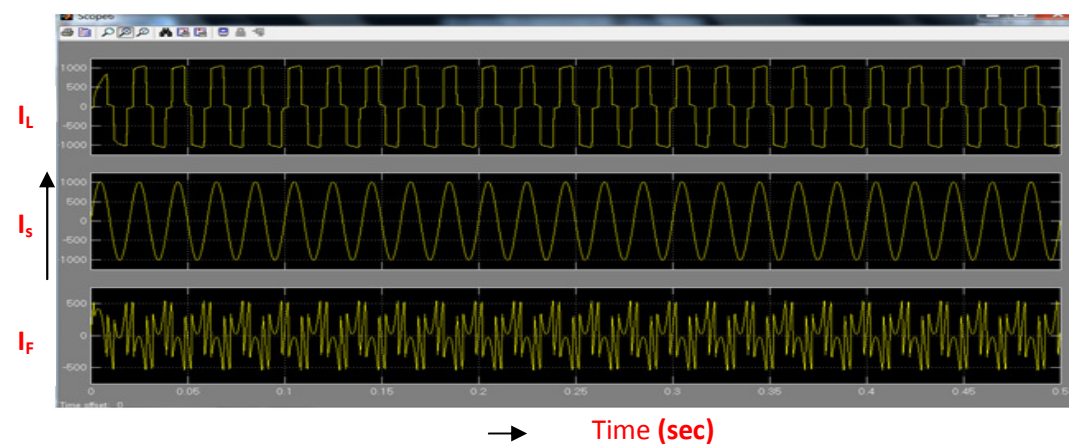


Fig.10. Wave form of Balanced load for HAPF-PSO Method

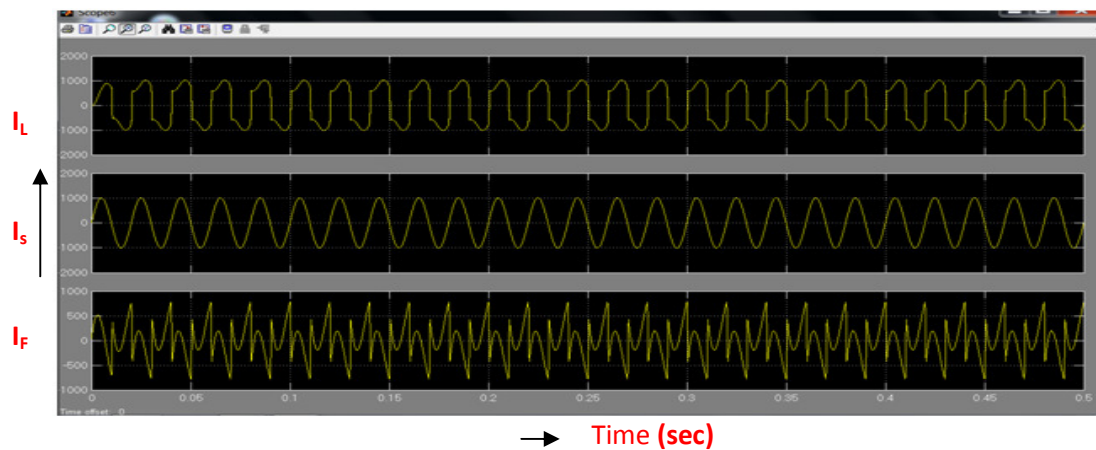


Fig.11. Wave form of Unbalanced load for HAPF-PSO Method

Table: VI. Results with Balanced Load

SCHEME	With PSO			With PSO		
	% THD	P.F	Reactive Power (VAR)	% THD	P.F	Reactive Power (VAR)
HAPF	0.47	0.9929	6.665	26.54	0.599	2.887

Table: VII. Results with Unbalanced Load

SCHEME	With PSO			With PSO		
	% THD	P.F	Reactive Power (VAR)	% THD	P.F	Reactive Power (VAR)
HAPF	0.49	0.9933	6.663	33.68	0.764	-8.0257

IV. CONCLUSION

The application of PSO to Hybrid active power filter is designed. The proposed control technique is found satisfactory to mitigate harmonics from utility current especially under balanced and unbalanced loading conditions. Thus, the resulting total current drawn from the ac mains is sinusoidal. The proposed design of SAF improves the overall control system performance over other conventional controller. The validity of the presented controllers was proved by simulation of a three phase four wire test system under balanced and unbalanced loading conditions. The proposed Hybrid shunt active filter compensate for balance and unbalanced nonlinear load currents, adapt itself to compensate variations in non linear load currents, and correct power factor of the supply side near to unity. Proposed APF topology limits THD percentage of source current under limits of IEEE-519 standard. It has also been observed that reactive power compensation has improved leading to power factor improvement with the PSO Technique.

REFERENCES

- [1] V. E. Wanger, "Effects of harmonics on equipment," *IEEE Trans. Power Del.*, vol. 8, no. 2, pp. 672–680, Apr.1993.
- [2] J. C. Das, "Passive filters-potentialities and limitations," *IEEE Trans. Ind.Appl.*, vol. 40, no. 1, pp. 232–241, Jan.Feb 2004.
- [3] K. K. Shyu, M. J. Yang, Y. M. Chen, and Y. F. Lin, "Model reference adaptive control design for a shunt active-power-filter system," *IEEE Trans. Ind. Electron.*, vol. 55, no. 1, pp. 97–106, Jan. 2008.
- [4] L. Asiminoaei, E. Aeloiza, P. N. Enjeti, and F. Blaabjerg, "Shunt activepower-filter topology based on parallel interleaved inverters," *IEEE Trans.Ind. Electron.*, vol. 55, no. 3, pp. 1175–1189, Mar. 2008.
- [5] A. Luo, Z. K. Shuai, W. J. Zhu, and Z. J. Shen, "Combined system for harmonic suppression and reactive power compensation," *IEEE Trans. Ind. Electron.*, vol. 56, no. 2, pp. 418–428, Feb. 2009.
- [6] B. Singh and V. Verma, "An indirect current control of hybrid power filter for varying loads," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 178–184, Jan. 2006.
- [7] D. Rivas, L. Moran, J. W. Dixon, and J. R. Espinoza, "Improving passive filter compensation performance with active techniques," *IEEE Trans. Ind. Electron.*, vol. 50, no. 1, pp. 161–170, Feb. 2003.
- [8] V. F. Corasaniti, M. B. Barbieri, P. L. Arnera, and M. I. Valla, "Hybrid active filter for reactive and harmonics compensation in a distribution network," *IEEE Trans. Ind. Electron.*, vol. 56, no. 3, pp. 670–677, Mar. 2009.
- [9] K. P. Lin, M. H. Lin, and T. P. Lin, "An advanced computer code for single-tuned harmonic filter design," *IEEE Trans. Ind. Appl.*, vol. 34, no. 4, pp. 640–643, Jul./Aug. 1998.
- [10] C. J. Chou, C. W. Liu, J. Y. Lee, and K. D. Lee, "Optimal planning oflarge passive-harmonic-filter set at high voltage level," *IEEE Trans. PowerSyst.*, vol. 15, no. 1, pp. 433–441, Feb. 2000.
- [11] Y. M. Chen, "Passive filter design using genetic algorithms," *IEEE Trans. Ind. Electron.*, vol. 50, no. 1, pp. 202–207, Feb. 2003.
- [12] Z. S. Guang, W. Y. Ping, and J. L. Cheng, "Adaptive genetic algorithm based optimal design approach for passive power filters," *Proc. Chin. Soc. Elect. Eng.*, vol. 24, no. 7, pp. 173–176, Jul. 2004.
- [13] Y. P. Chang and C. J. Wu, "Optimal multiobjective planning of large scale passive harmonic filters using hybrid differential evolution method considering parameter and loading uncertainty," *IEEE Trans. Power Del.*, vol. 20, no. 1, pp. 408–416, Jan. 2005.
- [14] B. Duro, V. S. Ramsden, and P. Muttik, "Minimization of active filter rating in high power hybrid filter system," in *Proc. IEEE Int. Conf. Power Electron. Drive Syst.*, Hong Kong, Jul. 1999, pp. 1043–1048.
- HE *et al.*: APPLICATION OF PSO TO PASSIVE AND HYBRID ACTIVE POWER FILTER DESIGN 2851
- [15] C. J. Ling, J. X. Jian, and Z. D. Qi, "Multi-object optimization of hybrid active power filter based on genetic algorithm," *J. Tsinghua Univ. Sci. Technol.*, vol. 46, no. 1, pp. 5–8, 2006.

- [16] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, Australia, Nov./Dec. 1995, vol. 4, pp. 1942–1948.
- [17] Y. D. Valle, G. K. Venayagamoorthy, S. Mohagheghi, J. C. Hernandez, and R. G. Harley, "Particle swarm optimization: Basic concepts, variants and applications in power systems," *IEEE Trans. Evol. Comput.*, vol. 12, no. 2, pp. 171–195, Apr. 2008.
- [18] L. S. Coelho and B. M. Herrera, "Fuzzy identification based on a chaotic particle swarm optimization approach applied to a nonlinear Yo-yo motion system," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 3234–3245, Dec. 2007.
- [19] S. Dutta and S. P. Singh, "Optimal rescheduling of generators for congestion management based on particle swarm optimization," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1560–1569, Nov. 2008.

AUTHORS

B. Suresh Kumar was born in Koppel, Parkas (Dot) India. He Received B.E from Bangalore University in 1999 and his M.Tech from JNT University of Hyderabad, India in 2003 both in Electrical and Electronics. He is at the finishing stage of PhD study in JNT University, Hyderabad, India. His employment experience included lecturer in the Department of Electrical and Electronics Engineering, At present He is working as Asst.professor in EEE dept CBIT Hyderabad, India. He is guiding 6 M.Tech Projects Power Quality. His special fields of interest are Power Quality, Power Systems, and Power Harmonics & Power Dynamics



K. Ramesh Reddy He obtained B.Tech from Andhra University in 1985, M.Tech from REC, Warangal in 1989 & Ph.D. from S.V. University, Tirupatl in 2004 India. He worked at GPREC, Kurnool as Teaching Assistant during 1985-1987. Also at KSRMCE, Kadapa as lecturer & Asst.Prof. From 1989 to 2000. During 2000-2003 he worked at LBRCE, Mylavaram as Professor & Head in EEE dept. Presently he is working as Professor & head EEE dept at G. Narayanamma Institute of Technology & Science, Hyderabad from 2003. He is also Dean of PG studies. He is having 22 years of teaching experience. He hold different positions as Chief Superintendent of Exams, ISTE chapter Secretary, IEEE branch Counselor, Officer in charge Library & Vice-Principal. He published 22 research papers in conferences & 6 papers in international Journals. He authored a textbook entitled "Modeling of Power System Components" published by Galgotia Publishers, New Delhi. He is reviewer for the international journal IEEE Transactions on Power Delivery and National Journal Institution of Engineers (India), Kolkata. Also he is a Technical committee member of IASTED, Calgary, Canada for conducting conferences at different countries. He is recipient of Best Engineering Teacher award by ISTE in 2007 & received Academic Excellence award from GNITS in 2008. At present he is guiding 9 Ph.D. students. His areas of research interest are Power Quality, Power Harmonics & Power Dynamics



S. Archana was born in Jaggayyapet, Krishna (Dt) India in 1988. She Received B.Tech degree in Electrical and Electronics Engineering from ANURAG Engineering College, JNT University, Hyderabad in 2009, Currently She is pursuing M.Tech-final year degree in Electrical and Electronics Engineering with the specialization of Power Systems and Power Electronics from CBIT, Hyderabad. Her special fields of interest are Power Quality, Power Systems & Power Electronics.



A SURVEY OF COUPLING MEASUREMENT IN OBJECT ORIENTED SYSTEMS

V. S. Bidve¹ and Akhil Khare²

¹Information Technology Department, M.Tech. (II), BVCOE, Pune, India

²Assistant Professor, Information Technology Department, BVCOE, Pune, India

ABSTRACT

Coupling measurement is a focus of study for many of the software professionals from last few years. Object-oriented programming is an efficient programming technique for programmers because of its features like reusability, data abstraction etc. Coupling is a very impotent factor in object-oriented programming for software quality measurement and used as predictors of software quality attributes such as fault proneness, impact analysis, ripple effects of changes, changeability etc. Many researchers have worked on coupling measurement and found various dimensions of coupling. Researchers have also worked on various aspects of coupling like static coupling measurement, dynamic coupling measurement, class level coupling, object level coupling etc. But still there is no standardization in the field of coupling measurement which is accepted worldwide. As a result of this it is very difficult to select any existing measure which obtain clear picture of state-of-art of coupling measurement for object-oriented systems. This paper analyses some terminologies of coupling measurement proposed earlier and discusses usefulness of each.

KEYWORDS: Coupling, measurement, object-oriented, dynamic, static

I. INTRODUCTION

Object oriented technology gaining significant importance in the field software development. To evaluate and maintain quality of object-oriented software there is a need to assess and analyse its design and implementation using appropriate measurement metrics [5]. A quality metrics should relate to external quality attributes of a design. External quality attributes include maintainability, reusability, error proneness, and understandability.

Based on observations and empirical studies, coupling has shown a direct impact on software quality [5]. In general, one of the goals of software designers is to keep the coupling in object-oriented system as low as possible. Classes of the system that are strongly coupled are most likely to be affected by changes and bugs from other classes. Such classes have more architectural importance; coupling measures helps in such happenings [2].

It is commonly observed in an object-oriented programming technique, inheritance and polymorphism is used more frequently. Static coupling measurement attributes are not sufficient to measure coupling due to inheritance and polymorphism.

As a result we focus on coupling measurement. We discuss various proposed dynamic coupling measurement metrics and their correlation quality attributes. We compare all measurement aspects and discuss in order to design uniform and standardized framework for coupling measurement.

The following section outlines the related work for object-oriented coupling metrics. Section 3 a detailed survey of existing coupling measures is carried out. In section 4 we provide comparative study of all the frameworks. Section 5 concludes the paper.

II. MOTIVATION

Object-oriented measurement has become popular area. There are large numbers of measures proposed for object oriented attributes such as coupling, inheritance, coherence. Also, there are several negative aspects regarding the manner in which the measures are developed and being developed. Coupling is a more complex software attribute in object oriented systems but our understanding about coupling measurement factors is poor. There is no standardization for expressing coupling measures; many measures are not operationally defined i.e. there is some ambiguity in their definitions. As a result, it is difficult to understand how different measures relate to one other and what their potential use is. All above aspects ultimately shapes to a need of detailed study of coupling measurement in object-oriented systems.

III. RELATED WORK

In this section we perform a detailed survey of existing coupling measurement frameworks in object-oriented systems.

3.1. Framework by Erik Arisholm [4]

The framework described by Erik considered following points regarding coupling measurement.

- Dynamic behavior of software can be precisely inferred from run time information.
- Static coupling measures may sometimes be inadequate when attempting to explain differences in changeability for object oriented design.

The author derived three dimensions of coupling.

1. Mapping : object or class
2. Direction: import or export
3. Strength: number of dynamic messages, distinct methods, or distinct classes.

The empirical evaluation of the proposed dynamic coupling measure consists of two parts.

- First to assess fundamental properties of the measure Second part evaluates whether the dynamic coupling measures can explain the change proneness of class.
- Erik used the concept of role-models for dynamic coupling measurement.
- Scenario: a specific sequence of interactions between the objects.
- Role: abstract representation of the functional responsibility of each object in a given scenario.

Object can have many roles because it may participate in many scenarios. The role-model reflects the dynamic coupling between the roles along three orthogonal dimensions: *direction*, *mapping* and *strength*.

- Direction of Coupling (Import and Export coupling): Dynamic import coupling counts the messages sent from a role, whereas dynamic export coupling counts the messages received.
- Mapping: Object-level and Class-level Coupling: *Object-level* coupling quantifies the extent to which messages are sent and received between the objects in the system. Dynamic, *class-level* coupling quantifies the extent of method dependencies between the classes implementing the methods of the caller object and the receiver object.
- Strength of Coupling: The strength of coupling quantifies the amount of association between the roles. It is of three types.
 1. *Number of dynamic messages*. Within a run-time session, to count the total number of times each message is sent from one role to another to implement a certain functional scenario.
 2. *Number of distinct method invocations*. To count the number of distinct method invocations between two roles.
 3. *Number of distinct classes*. To count the number of distinct classes.

Dynamic coupling is compared with static coupling and three important differences are Scope of Measurement, Dead code, Polymorphism. In all three measures dynamic coupling is considered more suitable than the static coupling. The relationship between dynamic coupling measures and the change proneness of the classes is explored by Erik and concluded that changes may prone to error.

3.2. Framework by R. Harrison, S. Counsell, R. Nithi [6]

This framework involves following points regarding coupling measurement.

1. Coupling between Object (CBO): Is a count of the number of classes to which a class is coupled. It counts each usage as a separate occurrence of coupling. This includes coupling via inheritance.
2. Number of Associations (NAS): is defined as the number of associations of each class. Counted from design documents. Counts repeated invocations as a single occurrence of coupling. This also includes coupling from inheritance.

Author considered that CBO is greater than NAS.

Three hypotheses related to coupling are investigated by the authors:

1. H1: As inter-class coupling increases, the understandability of a class decreases. This hypothesis is rejected by authors.
2. H2: As inter-class coupling increases, the total number of errors found in a class increases. This hypothesis is rejected by authors.
3. H3: As inter-class coupling increases, the error density of a class increases. This hypothesis is supported by authors.

To investigate these hypotheses author studied dependent variables such as

- Software Understandability (SU)
- Number of known errors(KE)
- Error per thousand non-comment source lines(KE/KNCSL)

Also, Coupling due to object as a parameter of methods and return type for a method is considered by authors.

3.3. Framework by Sherif M. Yacoub, Hany H. Ammar, and Tom Robinson [5]

The authors have referred many papers and conclude following points regarding coupling measurement.

Two design metrics are considered by authors.

1. Static: can only calculate design time properties of an application.
2. Dynamic: used to calculate actual runtime properties of an application.

Two types of coupling is considered by authors

1. Class level coupling (CLC): Only invocations from one method to another are considered.
2. Object level coupling (OLC): The invocations from one method to another and frequency of invocations at run time is also considered.

Authors also considered that, there is correlation between the number of faults and complexity of system. Therefore, static complexity is used to assess the quality of software. To measure dynamic complexity metrics authors used ROOM design charts. Cyclomatic complexity, operation complexity is calculated from ROOM charts.

Authors explained export and import object coupling with context, description, formula and impact on design quality attributes.

1. Export object coupling (EOC): Measure is a percentage of number of messages sent from object A to object B with respect to total number of messages exchanged during the execution of some scenario.
2. Import Object coupling (IOC): Measure is a percentage of the number of messages received by object A and was sent by object B *with* respect to the total number of messages exchanged during the execution of some scenario.

3.4. Framework by Erik Arisholm, Lionel C. Briand, and Audun Føyen [1]

The authors described many significant dynamic coupling measures and highlights way in which they differ from static measures. Authors collected the measures using UML diagrams and accounted precisely for inheritance, polymorphism and dynamic binding.

Classification of dynamic coupling measures

1. Entity of measurement: The entity of measurement may be a class or an object.
2. Granularity: The granularity can be class level or object level.
3. Scope: The objects and classes are to be accounted for measurement.

The authors captured situations for import and export coupling.

1. Dynamic messages: Total number of distinct messages sent from one object to other objects and vice versa, within the scope considered.
2. Distinct method invocations: Number of distinct methods invoked by each method in each object.
3. Distinct classes: Number of distinct classes that a method in a given object uses.

The measures described by authors are summarized into table 1.

Table 1. Heading and text fonts.

Direction	Entity	Strength
Import coupling	Class	Dynamic messages
		Distinct Methods
		Distinct classes
	Object	Dynamic messages
		Distinct Methods
		Distinct classes
Export coupling	Class	Dynamic messages
		Distinct Methods
		Distinct classes
	Object	Dynamic messages
		Distinct Methods
		Distinct classes

The authors described much more regarding polymorphism and dynamic binding using the above measures than the others and concluded that coupling is one of the factors which affect change proneness.

3.5. Framework by Lionel C. Briand, John W. Daly, and Jurgen Wust [3]

The authors identified six criteria of dynamic coupling measures.

1. The type of connection: What items (attribute, method or class) constitutes coupling.
2. The locus of impact: To decide whether to count import or export coupling.
3. Granularity of the measure: The level of detail at which information is gathered, i.e. what components are to be measured and how exactly the connections are counted.
4. Stability of server: There are two categories of class stability defined. The first is unstable classes which are subject to modification or development (user defined) and stable classes which are not subject o to change (library).
5. Direct or indirect coupling: To decide whether to count direct or indirect coupling. For example, if a method *m1* invokes a method *m2*, which in turn invokes a method *m3*, we can say that *m1* indirectly invokes *m3*. Methods *m1* and *m3* are indirectly connected.
6. Inheritance: inheritance-based vs. non-inheritance-based coupling, and how to account for polymorphism, and how to assign attributes and methods to classes.

IV. COMPARISON AND DISCUSSION OF EXISTING FRAMEWORK

A comparison shows that there are differences in the manner in which coupling is addressed. There are differences due to different points of focus by different authors. It is observers that there is no uniformity in the measurement. The significant differences are discussed in the following subsection.

4.1 Type of coupling

In most of the frameworks the entity of measurement is a class or an object. But, the mechanisms that constitute coupling are different. Erik uses the concept of role-model that constitutes dynamic coupling. Harrison considers coupling due to any means between the classes including parameter passing to a method, return type of a method. Sherif and his team consider invocation from one method to another as a coupling. Lionel and his team consider connection due attributes, classes, and

method as a coupling. There are differences in the mechanism that constitute coupling with respect each framework.

4.2 Strength of coupling

It depends on type of connection and frequency of connection between two classes. Different types of coupling have different strengths. Erik counts the strength in terms of number of dynamic messages, number of distinct method invocations, and number of distinct classes involved in coupling. Harrison counts it in term of number of classes to which a class is coupled and counted each invocation separately. Sherif and his team counts strength in terms of method invocations and frequency of invocation. Lionel and his team have considered granularity instead of strength. Measure for strength of coupling is not uniform it also varies as per author.

4.3 Direction of coupling

The framework by Erik distinguishes import and export coupling. Import coupling counts messages sent from a role, whereas export coupling counts the messages received. Harrison has not discussed anything regarding direction of coupling. Sherif and his team have explained import and export coupling with respect to total number of messages exchanged during scenario. Lionel and his team explained it as a locus of impact in which import coupling is analyzed as client whereas export coupling as server in their roles. Definition of import and export coupling is also ambiguous. There is need to clearly define the concept of client and server class.

4.4 Direct and indirect coupling

Only Lionel and his team have discussed the concept of direct and indirect coupling. The observation is that many measures stated have used the direct coupling but some of measures have used indirect coupling also. Consideration of direct or indirect measure is again a matter of discussion. Many authors have not defined direct and indirect coupling. There is a need to clearly define these terms and to show measures under the terms.

4.5 Stability of server class

This point is unique to the framework by Lionel and his team. Using a stable class is better than using an unstable class, because modifications which could ripple through the system are less likely to occur. The remaining frameworks have not discussed this point; this is again an important point. How stability of server class is important that is also not discussed by many authors.

4.6 Inheritance

Inheritance is very important aspect of dynamic coupling. It is observed that there is a need to consider inheritance based coupling in measurement. Erik has considered polymorphism as part of dynamic coupling but not discussed inheritance. Harrison has considered coupling due to inheritance but not given any measures of inheritance and non-inheritance based coupling. Sherif also has used ROOM charts which shows coupling due to inheritance but explicitly it is not discussed. Erik and his team accounted inheritance, polymorphism and dynamic binding using various levels of granularity. Lionel and his team have differentiated various measures under the category of inheritance based and non-inheritance based coupling. There is no clear picture of how to use inheritance in coupling. Every author has different idea regarding inheritance for coupling measurement.

4.7 Granularity

The granularity of the measure is the level of detail at which information is gathered. This is also an important point but not discussed by all the authors. Erik and Lionel have discussed the point but both have given a different explanation of the same. Very few authors have discussed this point. There is no clear understanding regarding granularity when we consider multilevel inheritance.

V. CONCLUSIONS

In this paper, we have studied five frameworks of dynamic coupling measurement for object-oriented systems. The motivation is to point out lack standardization and uniformity in the dynamic coupling

measurement. We have made comparison of all five frameworks with respect to total seven aspects. It is found that all frameworks differ in the definitions of measure, depth of measure, scope of measure and inclusion of points for coupling measurement. Many measures are ambiguous for e.g. type of coupling is an aspect in which cases which constitutes to a coupling are clearly not defined. Similarly, it is found with inheritance, strength of coupling and all other aspects of dynamic coupling measurement. Finally we come to the conclusion with following points.

- There is need of standardization in the field of dynamic coupling measurement
- Clear definition of every aspect of a measurement is needed
- Scope of measurement is needed to define for each measure
- Every measure must be empirically supported

These are problems we faced in the study of the various frameworks emerged as ideas to design a new framework model for dynamic coupling measurement.

REFERENCES

- [1] Erik Arisholm, Lionel C. Briand, and Audun Føyen, "Dynamic Coupling Measurement for Object-Oriented Software," *IEEE Transactions on Software Engineering*, 30(8), 2004.
- [2] Denys Poshyvanyk, Andrian Marcus, "The Conceptual Coupling Metrics for Object-Oriented Systems," *ICSM '06 Proceedings of the 22nd IEEE International Conference on Software Maintenance*, 2006.
- [3] Lionel C. Briand, John W. Daly, and Jürgen Wüst, "A unified framework for coupling measurement in object-oriented systems," *IEEE Transactions on Software Engineering*, 25(1), 91-121, 2002.
- [4] Erik Arisholm, "Dynamic Coupling Measures for Object-Oriented Software," *IEEE Symposium on Software Metrics in Proc.8*, 33-42, 2002.
- [5] Sherif M. Yacoub, Hany H. Ammar, and Tom Robinson, "Dynamic Metrics for Object Oriented Designs," *Software Metrics Symposium, Proceedings. 6*, 50-61, 2002.
- [6] R. Harrison, S. Counsell, R. Nithi, "Dynamic Metrics for Object Oriented Designs," *Software Metrics Symposium, Proceedings. 5*, 150-157, 2002.
- [7] S.R. Chidamber, C.F. Kemerer, "Towards a Metrics Suite for Object Oriented design", in A. Paepcke, (ed.) *Proc. Conference on Object-Oriented Programming: Systems, Languages and Applications(OOPSLA'91)*, October 1991. Published in *SIGPLAN Notices*, 26 (11), 197-211, 1991.
- [8] S.R. Chidamber, C.F. Kemerer, "A Metrics Suite for Object Oriented Design", *IEEE Transactions on Software Engineering*, 20 (6), 476-493, 1994.
- [9] V. R. Basili, L. C. Briand, and W. L. Melo. A validation of object-oriented design metrics as quality indicators. *IEEE Transactions on Software Engineering*, 22(10):751{761, 1996.
- [10] F. Abreu, M. Goulão, R. Esteves, "Toward the Design Quality Evaluation of Object-Oriented Software Systems", *5th International Conference on Software Quality*, Austin, Texas, USA, October 1995.
- [11] V. Basili, L. Briand, W. Melo, "Measuring the Impact of Reuse on Quality and Productivity in Object-Oriented systems", *Technical Report*, University of Maryland, Department of Computer Science, CSTR-3395, January 1995.
- [12] V.R. Basili, L.C. Briand, W.L. Melo, "A Validation of Object-Oriented Design Metrics as Quality Indicators", *IEEE Transactions on Software Engineering*, 22 (10), 751-761, 1996.
- [13] L. Briand, P. Devanbu, W. Melo, "An Investigation into Coupling Measures for C++", *Technical ReportISERN 96-08, IEEE ICSE '97*, Boston, USA, (to be published) May 1997.
- [14] L. Briand, K. El Emam, S. Morasca, "Theoretical and Empirical Validation of Software Product Measures", *Technical Report*, Centre de Recherche Informatique de Montréal, 1995.
- [15] L. Briand, S. Morasca, V. Basili, "Measuring and Assessing Maintainability at the End of High-Level Design", *IEEE Conference on Software Maintenance*, Montreal, Canada, September 1993.
- [16] L. Briand, S. Morasca, V. Basili, "Defining and Validating High-Level Design Metrics", *Technical Report*, University of Maryland, CS-TR 3301, 1994.
- [17] L. Briand, S. Morasca, V. Basili, "Property-Based Software Engineering Measurement", *IEEE Transactionsof Software Engineering*, 22 (1), 68-86, 1996.
- [18] S.R. Chidamber, C.F. Kemerer, "Towards a Metrics Suite for Object Oriented design", in A. Paepcke, (ed.) *Proc. Conference on Object-Oriented Programming: Systems, Languages and Applications (OOPSLA'91)*, October 1991. Published in *SIGPLAN Notices*, 26 (11), 197-211, 1991.
- [19] S.R. Chidamber, C.F. Kemerer, "A Metrics Suite for Object Oriented Design", *IEEE Transactions onSoftware Engineering*, 20 (6), 476-493, 1994.

- [20] N.I. Churcher, M.J. Shepperd, "Comments on 'A Metrics Suite for Object-Oriented Design'", *IEEE Transactions on Software Engineering*, 21 (3), 263-265, 1995.
- [21] N.I. Churcher, M.J. Shepperd, "Towards a Conceptual Framework for Object Oriented Software Metrics", *Software Engineering Notes*, 20 (2), 69-76, 1995.
- [22] P. Coad, E. Yourdon, "Object-Oriented Analysis", *Prentice Hall*, second edition, 1991.
- [23] P. Coad, E. Yourdon, "Object-Oriented Design", *Prentice Hall*, first edition, 1991.
- [24] J. Eder, G. Kappel, M. Schrefl, "Coupling and Cohesion in Object-Oriented Systems", *Technical Report*, University of Klagenfurt, 1994.
- [25] N. Fenton, "Software Metrics: A Rigorous Approach", *Chapman and Hall*, 1991.
- [26] M. Hitz, B. Montazeri, "Measuring Coupling and Cohesion in Object-Oriented systems", in *Proc. Int. Symposium on Applied Corporate Computing*, Monterrey, Mexico, October 1995.
- [27] M. Hitz, B. Montazeri, "Chidamber & Kemerer's Metrics Suite: A Measurement Theory Perspective", *IEEE Transactions on Software Engineering*, 22 (4), 276-270, 1996.
- [28] I. Jacobson, M. Christerson, P. Jonsson, G. Overgaard, "Object-Oriented Software Engineering: A Use Case Driven Approach", *ACM Press/Addison-Wesley*, Reading, MA, 1992.
- [29] E. Arisholm, "Empirical Assessment of Changeability in Object-Oriented Software," PhD Thesis, Dept. of Informatics, Univ. of Oslo, ISSN 1510-7710, 2001.
- [30] Oslo, ISSN 1510-7710, 2001.
- [31] E. Arisholm, "Dynamic Coupling Measures for Object-Oriented Software," *Proc. Eighth IEEE Symp. Software Metrics (METRICS '02)*, pp. 33-42, 2002.
- [32] E. Arisholm, D.I.K. Sjøberg, and M. Jørgensen, "Assessing the Changeability of Two Object-Oriented Design Alternatives—
- [33] A Controlled Experiment," *Empirical Software Eng.*, vol. 6, no. 3, pp. 231-277, 2001.
- [34] E. Arisholm, L.C. Briand, and A. Føyen, "Dynamic Coupling Measurement for Object-Oriented Software," *Technical Report*
- [35] 2003-05, Simula Research Laboratory, <http://www.simula.no/~erika>, 2003.
- [36] G. Booch, J. Rumbaugh, and I. Jacobson, *The Unified Modeling Language Users Guide*. Addison-Wesley, 1998.
- [37] L. Bratthall, E. Arisholm, and M. Jørgensen, "Program Understanding Behaviour During Estimation of Enhancement Effort on Small Java Programs," *Proc. Third Int'l Conf. Product Focused Software Process Improvement (PROFES 2001)*, 2001.
- [38] L.C. Briand and J. Wuest, "Empirical Studies of Quality Models in Object-Oriented Systems," *Advances in Computers*, vol. 59, pp. 97-166, 2002.
- [39] L.C. Briand and Y. Labiche, "A UML-Based Approach to System Testing," *Software and Systems Modeling*, vol. 1, no. 1, pp. 10-42, 2002.
- [40] L.C. Briand, J. Daly, and J. Wust, "A Unified Framework for Cohesion Measurement in Object-Oriented Systems," *Empirical Software Eng.*, vol. 3, no. 1, pp. 65-117, 1998.
- [41] L.C. Briand, J.W. Daly, and J. Wust, "A Unified Framework for Coupling Measurement in Object-Oriented Systems," *IEEE Trans. Software Eng.*, vol. 25, no. 1, pp. 91-121, 1999.
- [42] L.C. Briand, J. Wust, and H. Lounis, "Using Coupling Measurement for Impact Analysis in Object-Oriented Systems," *Proc. Int'l Conf. Software Maintenance (ICSM '99)*, pp. 475-482, 1999.
- [43] F. BritoeAbreu, "The MOOD Metrics Set," *Proc. ECOOP '95 Workshop Metrics*, 1995.
- [44] M. Cartwright and M. Shepperd, "An Empirical Investigation of an Object-Oriented Software System," *IEEE Trans. Software Systems*, vol. 26, no. 8, pp. 786-796, 2000.
- [45] Systems, vol. 26, no. 8, pp. 786-796, 2000.
- [46] M.A. Chaumon, H. Kabaili, R.K. Keller, F. Lustman, and G. Saint-Denis, "Design Properties and Object-Oriented Software Changeability," *Proc. Fourth Euromicro Working Conf. Software Maintenance and Reeng.*, pp. 45-54, 2000.
- [47] S.R. Chidamber and C.F. Kemerer, "A Metrics Suite for Object-Oriented Design," *IEEE Trans. Software Eng.*, vol. 20, no. 6, pp. 476-493, 1994.
- [48] S.R. Chidamber, D.P. Darcy, and C.F. Kemerer, "Managerial Use of Metrics for Object-Oriented Software: An Exploratory Analysis," *IEEE Trans. Software Eng.*, vol. 24, no. 8, pp. 629-637, 1998.
- [49] I.S. Deligiannis, M. Shepperd, S. Webster, and M. Roumeliotis, "A Review of Experimental Investigations into Object-Oriented
- [50] Technology," *Empirical Software Eng.*, vol. 7, no. 3, pp. 193-232, 2002.
- [51] G. Duntelman, *Principal Component Analysis*. SAGE, 1989.
- [52] K. El Emam, S. Benlarbi, N. Goel, and S.N. Rai, "The Confounding Effect of Class Size on the Validity of Object-Oriented Metrics,"

- [53] IEEE Trans. Software Eng., vol. 27, no. 7, pp. 630-650, 2001.
- [54] R.J. Freund and W.J. Wilson, Regression Analysis: Statistical Modeling of a Response Variable. Academic Press, 1998.
- [55] Jakarta, "The Apache Jakarta Project," <http://jakarta.apache.org/>, 2003.
- [56] Java.net, "Java Compiler Compiler (JavaCC)," <https://javacc.dev.java.net/>, 2003.
- [57] H. Kabaili, R. Keller, and F. Lustman, "Cohesion as Changeability Indicator in Object-Oriented Systems," Proc. IEEE Conf. Software Maintenance and Reeng. (CSRM), pp. 39-46, 2001.
- [58] A. Lakhota and J.-C. Deprez, "Restructuring Functions with Low Cohesion," Proc. IEEE Working Conf. Reverse Eng. (WCRE), pp. 36-46, 1999.
- [59] G. Myers, Software Reliability: Principles and Practices. Wiley, 1976.
- [60] H. Sneed and A. Merey, "Automated Software Quality Assurance," IEEE Trans. Software Eng., vol. 11, no. 9, pp. 909-916, 1985.
- [61] M.M.T. Thwin and T.-S. Quah, "Application of Neural Networks for Software Quality Prediction Using Object-Oriented Metrics," Proc. IEEE Int'l Conf. Software Maintenance (ICSM), 2003.

Authors

V. S. Bidve: Computer engineering from University of Aurangabad and the M. Tech pursuing from BVUCOE, Pune. He has nine years of teaching experience in Pune and Mumbai. He is now working as a lecturer in the Department of information technology SKNCOE, Pune.



A. R. Khare: Has completed bachelor degree in computer engineering from Bhopal University, India and M. Tech. from same University. Pursuing P. hD. In the field of computer engineering. Working as Assistance Professor in Information Technology Department of BVCOE, Pune. Having 10+ years of teaching experience. Working as a PG coordinator for IT department and guiding number of students for their project work and various academic activities.



THE COMPUTER ASSISTED EDUCATION AND ITS EFFECTS ON THE ACADEMIC SUCCESS OF STUDENTS IN THE LIGHTING TECHNIQUE AND INDOOR INSTALLATION PROJECT COURSE

İsmail Kayri¹, Muhsin Tunay Gençoglu² and Murat Kayri³

¹Faculty of Technical Education, Electrical Department, Batman University, Batman, Turkey

²Faculty of Engineering, Electrical Department, Firat University, Elazığ, Turkey

³Department of Computer & Instructional Technology, Yuzuncu Yil University, Van, Turkey

ABSTRACT

The purpose of this study is to investigate the effects on students' academic success of a visual didactic material developed in a computing environment which is believed it will enlighten the students during the process of completing the project and the explanation of "the Lighting Technique and Internal Installation Project" course that is taught in the curriculum of the electrical and electronic parts from institutions who are given formal and widely education such as Technical Education and Engineering Faculties, Vocational Colleges, Public Education Centers, Industrial Vocational High schools which are the backbone of vocational and technical education. In addition, the use of the educational software as in this area as a didactic material that is developed for the mentioned course is determined as a subsequent goal. To test the effectiveness of the developed educational software in the learning process there are two measurement tools for the cognitive dimension developed according experts and accordant findings of these measurement tools the effectiveness is examined.

KEYWORDS: *Computer-assisted teaching, Computer-assisted education, Electrical Installation Project, Visual Education*

I. INTRODUCTION

Vocational and technical education can be defined as "the whole teaching, supervision, management with coordination, organization, development, research and planning activities of any kind vocational and technical training in the industry, agriculture and service sectors within the integrity of the national education" [1].

In developed western countries is vocational training defined as a vocation branch that aims to gain a career through practical activities or manual skills [2]. The purpose of vocational-technical education is in generally to educate and train individuals as qualified work force for employment in the industry, trade and service sectors, and to give the basic education that is required for the transition to higher education institutions which is the continuation of their vocation [3].

During examination of the curricula from developed countries and European countries in particular, there is a presence of a linear proportionality observed between their development and their importance for vocational and technical training.

The self-sufficiency of the countries at the point of producing goods and services and getting healthy along with other countries with which they have economic relations, using the current technology and producing new technologies is related with the importance they show for vocational and technical

training in their development plans. Therefore, instructors strive to make the content more effectively when programming the learning processes.

An effective education can be achieved by discussing and eliminating time, space and economic concerns which limit the instructors as well as the students. This discussion encloses a wide area that includes the systematic and reflective transferring process of the teaching design on the teaching and learning principles, didactic materials, didactic activities, information sources and evaluation plans [4].

While these concepts are discussed, the learning of the student by internalizing the given information comes to the fore in the 'teacher, student and environment triangle' which has an important role in the teaching process. By thinking that each learner has different psychological, social and cognitive development characteristics, the importance remains of converting the teacher-centered education into student-centered and preparing in this context the teaching desing by taken the learning style of the learner into account [5].

According research [6,7], it is determined that taking the learner to the center of the learning process and the education processes that are established by considering the learning methods ameliorate the creativity, motivation and the academic achievement. It is putting forth that in an education designed by including the learning methods into the learning process, the learner can use the information more effective by remembering it for a longer time [8]. When on the other hand studies on learning styles are examined [9,10,11], it is observed that as a result of teaching designs that are configured by the participation of the learner in the learner process and that are supported with technologies, the academic success and performance increased and that a more positive attitude to learning is developed. The concerned findings of the research puts forth that the teacher need to prepare the learning environment; method, technique and didactic materials according the properties of the student and the lesson [5]. Crowded classrooms, unmet education demands, facilities, equipment inability, unbalanced distribution in terms of equality of opportunity, unmet individual needs, yield losses of students' success and similar problems are considered as the crucial characteristic problems of the traditional education systems. The Turkish vocational and technical education remained behind the developed countries in terms of the number students per teaching staff. The number of student per teaching staff ranges in the four-year vocational and technical training faculties from 22,7 to 33,6, in the Vocational School of Higher Education it is 60,8 and in the secondary school it is 31,7. In developed countries these numbers mainly ranges from 5 to 10. The most important problems of the vocational and technical education institutions in addition to the problems around the number of student and lack of teaching staff are infrastructure, technological equipment and the deficit of laboratory and workshop [12].

The sheer number of students per teaching staff will negatively affect the learning process in classes where computers are used as tool. The elimination of these problems will be possible with bringing the invididual to the fore during the education process, with a student-centered education, by designing, applying, evaluating and developing the techniques to be applied during the process, with a contemporary understanding and in accordance with the needs of the time [13].

The each day growing complexity of education, the rise of the information to learn, the need for qualified and modern education require the use of computers as a tool in education. The use of technology in education will provide that the education will be carried out in accordance with the needs of the era as well as that the highest appropriate yield will be received from education [14]. The computer which is one of the technological capabilities and a basic element of culture in our century, has become a tool which use is rapidly spread [15].

Computer-assisted education is the set of applications about the use of the computer as a tool for a directly offer of the course contents, for repeating knowledge gained in other ways, for solving problems, for practice and for similar activities [15]. The education technology that will meet the specified requirements includes the targeted application of human power and other resources so learning can be convert to easy, concrete, rich, meaningful, motivational, encouraging, efficient and qualified activities by optimizing the teaching-learning processes [16].

In this study it is aimed to present the above-mentioned advantages of educational technology in a particular course. The research model has an experimental – control group pattern, and the work group consist the fourth grade students who follow the Lighting Technique and Internal Installation Project course of the Electricity Teacher Training Department from the Faculty of Technical

Education which is located in the Southeast Anatolia Region. The experimental and control group are obtained by randomly splitting 30 students of the same branch into equal groups.

The nature of the executed course includes a project-based process for the experimental group as well as for the control group. After the theoretical dimension of the course was explained to the students during five weeks in the same environment and equal time frames, the students in the control group received a single purpose educational software CD that is developed by researchers. The students of the control group were released to take advantage of any kind of material including internet after the theoretical process of the instructor. Both groups of students were asked to project the strong and weak current installation on their own architectural projects. This is the common method throughout the country to train students in this course. There has been no intervention in the experimental-control groups.

1.1 Educational Technology

Educational Technology is the functional structuration of the learning or education processes by running knowledge and skills to dominate education in general and learning in particular.

According to another definition, educational technology is a branch of science that studies the ways to carry individuals to the special purposes of education by using wisely and skillfully the accessible human power and other sources related to education and based on relevant data about communication and learning from behavioral sciences with appropriate methods and techniques and by evaluating the results.

The current sense of educational technology is a technology related to education sciences that develops, applies, evaluates and manages appropriate designs by running the whole concerning elements (human power, knowledge, method and techniques, tools and necessary arrangements) for a systematic and scientific analysis of all aspects of the phenomenon of human learning and to bring solutions. In other words, educational technology is an original discipline about learning and teaching processes [17].

The reflections of the effective use of education technologies on the learning process can be listed as follows:

Technologies:

- improves the quality of learning.
- reduces the time spent to reach the goal of the students and teachers.
- improves the effectiveness of teaching.
- reduces the cost of education without reducing the quality.
- involves the student in the environment [18].

Some of the facilities of the modern educational technology provided for educational applications can be listed as follows:

- Providing freedom and initiative,
- Enlarging options,
- Saving the individual from the monopoly of the group,
- Providing the student the opportunity of individual and independent learning,
- Providing information of the first source,
- Solving the inequality of opportunities,
- Providing quality in education,
- Providing standardization, diversity and flexibility in education programs,
- Increasing the speed of learning,
- Adding at the same time individuation and popularization properties to the educational services,
- Providing the opportunity to increase the efficiency and effectiveness of the learning-teaching processes [17].

1.2 Computer-Assisted Teaching

Thanks to the features like quickly processing, saving and delivering information, the computer has become the most wanted tool in education. The use of an intensive technology is indeed regarded as strange because of the human labor in the activities of measuring and evaluating the success, in the

student guidance-consultancy work and in the running of educational services that have become complicated depending on the growing number of student in researchs on education. The intensive usage of technological resources in education has received a wide acceptance and the applications has increased. Therefore the apply in education of computers which are used in every stage of life can not be opposed. Computers are filling an important gap in several topics in which the tools and materials used in traditional education are insufficient. Many matters that are difficult or impossible to perform in traditional education can be accomplished with computers.

In computer-assisted education, the computer has an application area where it can be used supportive, with the teacher or individually or with other methods and techniques. Therefore, the computer-assisted education is seen as the most promising between the methods in education services. It is indicated that using a virtual lab in engineering education has a positive impact on factors such as the involvement of students in class, the self-confidence and motivation [19], provides students an individual learning environment [20], gives the students the opportunity to gain a wide variety of experiences about different approaches and helps to learn in an interactive and meaningful way [21].

The computer-assisted education as an educational environment where the teacher prepares the learning environment, recognizes the students abilities, accomplishes the activities like repeating, practising, directing and personalization according to the students' capabilities requires the usage of the computer in different place, time and ways according to the learning objectives which are determined in agreement with the structure of the teaching matter [22].

The following findings were obtained in some international researchs about the usage of the computer in education;

1. The computer helps the students to achieve their academic goals.
2. Compared to the traditional education, computer programs provide a saving of the learning time between 20% and 40%.
3. Using the computer in the field of education has a positive impact on the students' success and increases the motivation.
4. The effectiveness of educational software is playing an important role in the success of the computer-assisted education [23].

1.2.1 Benefits of Computer-Assisted Teaching

The benefits of computer-assisted teaching can be listed as followed;

The materials that were not understood can be repeated several times by the students.

There is no dependency on someone else and each student learns at own pace.

During the implementation of the computer-assisted training the student must participate actively in class.

- Errors and shortages will be discovered and corrected while learning
- The student has always a chance to answer again.
- It keeps the students' interest to the class always alive.
- Gives the teacher the opportunity to deal more closely with the students by saving them from works like repeating and correction etc.
- Dangerous and expensive studies can be easily simulated with computer-assisted teaching
- The students can learn more quickly and in a systematic way.
- The level of students' attention can be held very high through drawings, colors, figures, and images the students see while following the class.
- Learning is demoted into small units so that success can be achieved step by step by testing it on these units [22].

In addition, according to performed research [24], the most import benefit of the use of computers for education is the facilitation of the access to information.

II. METHOD / PROCEDURE

2.1 The Goal and Importance of Research

The current teaching in our education institutions are not going forward than an activity aimed to memorise rules by using the blackboard and the textbook which is a teacher-centered education. In

addition to the didactic methods applied in the existing system there is need for taking advantage of computer-assisted training applications like demonstration, simulation, practice and exercise, establishing a dialogue, problem solving, didactic games, information store, creative activities and testing [25]. The declines in computer prices and the need for computer in the communication and information age made the computer indispensable for especially new generation people. The disadvantages faced in education institutions like the insufficiency of time, place, technology and teaching staff created the need for learning activities outside of the school. At this point the widespread of computers that are present in almost every household has played an important role. University students, high school and even elementary school students are using computers to reinforce the treated material at school.

While in many branches of science the effects of the computer-assisted education on the motivation of students and their academic success are examined, there is no found any research in electrical and electronical science branch in literature containing the project of a building's electrical installation.

In this research is the following subject examined: The effects on students' academic success from the software that is developed to support the traditional education method in the Electrical Installation Project course which is an important part of the Electrical and Electronical Science branch.

2.2 Work Group

The workgroup of this research is formed by the students of the fourth class who follow the Lighting Technique and Indoor Installation Project course in the Electricity Teacher Training department of the Faculty of Technical Education of the University of Batman. The experimental group is obtained by randomly splitting 30 students studying at the same branch into 2 groups of 15 students. There is no significant difference identified between the grade point averages of the groups belonging to earlier periods. All members of the group are male.

The members of group A and B received with a projection device at the same time periods and by the same instructor the required theoretical knowledge to project the strong and low current installation of a building in an AutoCAD software environment. Both groups received resources which contain the theoretical information about the course. The members of group B received also an audio-visual education CD that is developed by researchers through various programs in a computer environment. This education CD contains an audio-visual description in an AutoCAD drawing program environment of how to achieve all required processes from the creation of the symbols in the program environment that will be used in the electrical installation's project of a building, as well as the drawing of the strong and low current installation from the same building to the required calculations. This CD includes also the theoretical materials that are sorted systematically so that students can easily access when they require it. After that the groups of students are equipped with theoretical knowledge, they were asked to project the strong and low current installation on an architectural plan provided by the instructor or by the students.

2.3 Private Educational Software

High-tech products such as computers, television and internet are used in education area to support the training. Many educational institutions are choosing to develop new alternatives to take advantage of the benefits provided by new technologies and to improve the usability of their current education programs. It is intended to give more people an education outside of the traditional education approach by using this new methods than to these alternatives [13]. The private didactic software which is prepared in a computer environment contains the audio-visual media that describes on a sample architectural plan the required processes for the project of an electrical installation of a building. With this software the students can repeat the issues that remain limited with the commentaries of the instructor in the classroom. Again, thank to the software including CD, they can see the path to follow to achieve the required calculation and the drawing of the electrical project.

In the drawing of the project the AutoCAD program were preferred. Although the students master the principles of project drawing some students have had problems on the use of the program and they have transferred these problems to the instructor during the class as well as outside the classroom. This leads to a loss of time in the educational process. The explanation in the private didactic software is mostly done through the AutoCAD program. Thus the students can watch infinitely the processes so one assumes that the students will ask less questions to the instructors. The required operations to

make the necessary calculations and to draw the project are systematically determined so that the resolution of the problems faced by the students can be reached quickly.

The prepared private didactic software also contains regulations for the drawing of the electrical installation project whereby the students possess an offline source about legislation. In short, the students who will project the electrical installation of a building can reach all required operations to execute and complete the project independently from space and time thanks to these private educational software.

The videos which constitute the most important leg of a private didactic software were prepared by the Screen Recorder module of the Blueberry Flashback Professional software. The video comments that were created with the same module are technically optimized in terms of quality and performance. Audio files that were mounted on some videos are recorded with the Polderbits Sound Recorder and Editor software and the essential quality optimization were done. The image files used in the software were optimized for use with the Adobe Photoshop editor. All of these media has been turned into an autorun CD within in a suitable menu structure and duplicated through the Camtasia Studio software.

Before the prepared private didactic software were applied on the experimental group, it is applied by researchers in the computer lab to detect and eliminate unforeseen technical problems by getting the views on the software of 30 final year students who are studying Computer and Instructional Technologies at the Education Faculty of the 'Yüzüncü Yıl' University. The percentage and frequency results concerning the obtained views are presented in Table 1.

Table 1. View of students about the Private Educational Software

CRITERIA	Very bad		Bad		Mediocre		Good		Very Good	
	f	%	f	%	f	%	f	%	f	%
N=30										
1. The level of attractiveness of video and animations	-	-	-	-	9	30	15	50	6	20
2. Easy usage of interfaces	-	-	-	-	4	13,3	16	53,3	10	33,3
3. Understandability of the content	-	-	-	-	9	30	14	46,6	7	23,3
4. Systematic transitions between topics	-	-	-	-	7	23,3	18	60	5	16,6
5. Color matching between text and graphics	-	-	-	-	5	16,6	15	50	10	33,3
6. Functionality of the transition buttons	-	-	-	-	7	23,3	20	66,6	3	10
7. Density of the graphics display	-	-	-	-	6	20	15	50	9	30
8. Readability of the screen	-	-	-	-	6	20	19	63,3	5	16,6
9. Flexibility of the video playback buttons	-	-	-	-	4	13,3	11	36,6	15	50
10. The sound quality and adjustability	-	-	-	-	16	53,3	14	46,6	-	-
11. Loading and execution speeds of videos	-	-	-	-	15	50	12	40	3	10
12. Suitability of the font and point values of the characters	-	-	-	-	7	23,3	13	43,3	10	33,3

By analyse Table 1 we can see that the students' views about the articles 10 and 11 are middle and about the other articles it is good or very good.

The reason of the high frequency at the middle selection of students' views about the article concerning the sound quality and adjustability is that 22050 Hz-48 Kbs-Mono were selected to provide that the audio files take a small place. These selections were changed as 44100 Hz-64 Kbs-Mono and converted into a Compact Disc quality. The reason that the students' view on the article concerning the load and execution speeds of videos are concentrated at the middle selection is that the graphic intensity is kepted at a high level to obtain a clear image.

The different hardware values of the computers used by the students produced such a result. Thinking about the hardware sufficiency of the computers belonging to the members of the group that will receive a CD; the resolution, frequency and color quality are reduced from 800x600-70 Hz-32 Bit to 800x600-60 Hz-16 so that the negativity stipulated in article 11 will disappear.

2.3.1 Operation principle of the Private Educational Software

The private didactic software developed by researchers through a variety of media development software is transformed into an autorun CD. After the installation of the CD on the computer, the main page of the software appears on the screen (Figure 1).

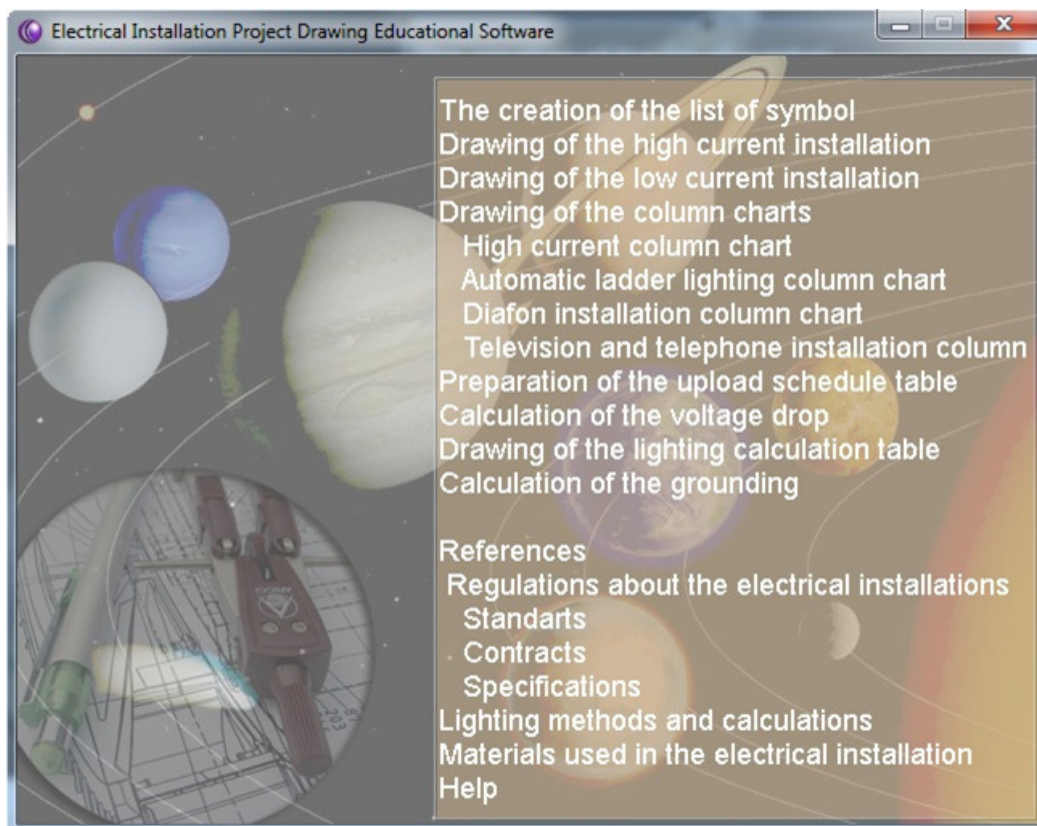


Figure 1. Main page of the Electrical Installation Project Drawing Educational Software

The main page of the software consists 3 parts:

1. Visual and audio videos commentary: This part contains the visual and audio videos on the architectural project prepared by researchers. The students can get support on concerning issues in the process of creating the project.
2. Referenced documents: This section contains the offline sources which includes the principles of the Electrical Installation Project drawing developed by researchers. It is assumed that the students can faster attain the desired information thanks to this feature.
3. Help: This part contains the help topics that include the operation principle of the program for an effective use of the private educational software by students.

The crucial part of the software is formed by audio and videos developed by researchers that systematically demonstrate from start to finish the project of an electrical installation on an architectural plan (Figure 2). Thanks to these videos, students can resolve the problems they face in the executing process of their own projects. In this way, the need to communicate with the instructor for every encountered problem will be eliminated.

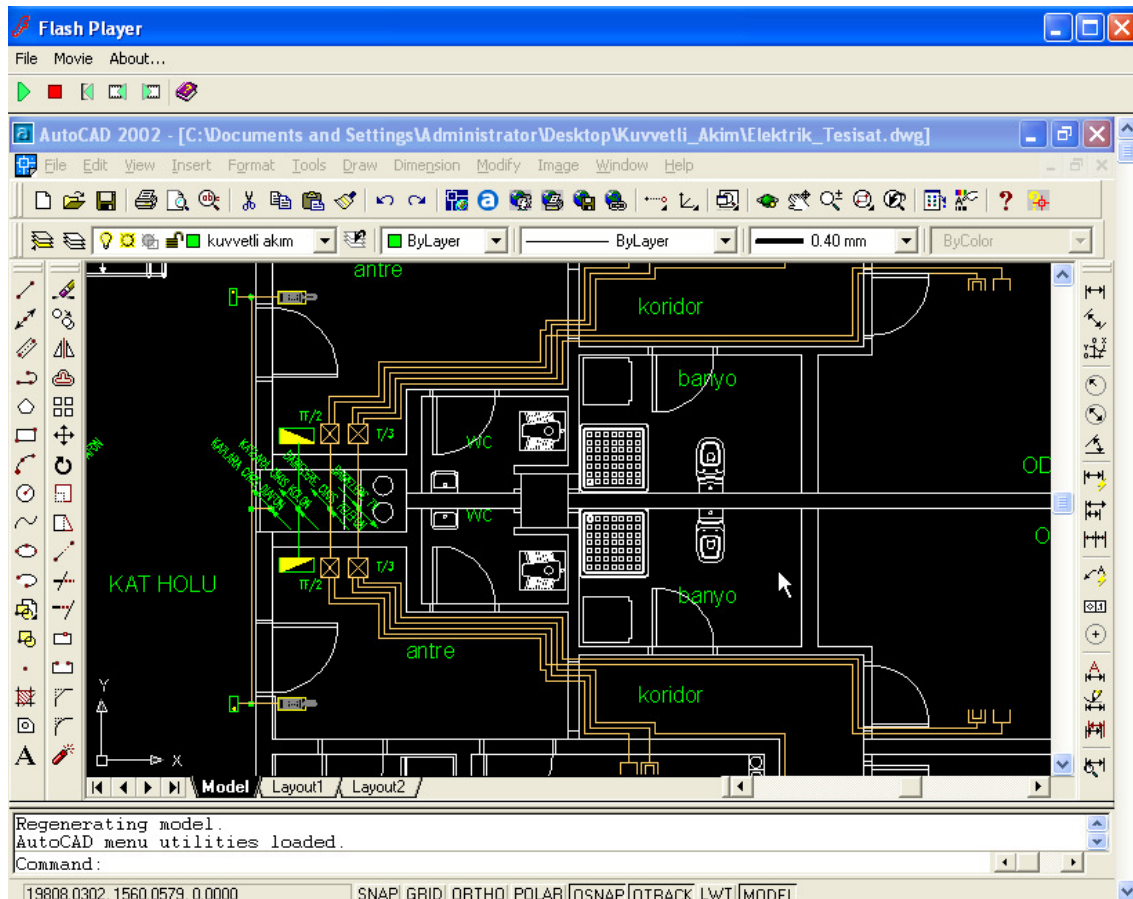


Figure 2. A screenshot of a video from the Educational software

III. FINDINGS AND COMMENTARY

After the transfer of the necessary theoretical knowledge by the traditional education system on equal terms and times to the groups A and B that constitute the research groups, were exercises made on sample projects with a projection device. The students of group A and B began to undertake their projects after the suitability of the students' architectural projects has been confirmed by the instructors. The developments about the conduct of the students' projects were assessed through scales prepared by researchers during a process of approximately 30 days (Table 2).

In the scale showing that the students in group B are in general more successful the remarkable items of success are as followed:

- The accuracy of drawn symbols to be used in the Project
- Compliance of drawing elements with the regulation places
- Conformity of the socket number and forces with regulation
- Adequacy of management in the drawing layer
- Accuracy of calculation of the voltage drop
- Accuracy of the upload table
- Accuracy of calculation of the grounding
- The correct use of time in the project execution process

The research shows that the prepared training Cd assisted the students in the required calculations and drawings of the project. Item 13 of the scale especially shows that the students who utilized the training Cd did used the time correctly, while the students in group A couldn't use it correctly because they often felt the need to consult the educational staff.

Table 2. Findings concerning the undertake of the project from the group member

CRITERIA	Students group A										Students group B									
	Weak		Passes		Mediocre		Good		Very Good		Weak		Passes		Mediocr e		Good		Very Good	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%
N=15																				
1.The accuracy of drawn symbols to be used in the project	1	6,6	3	20	7	46,6	4	26,6	-	-	-	-	1	6,6	2	13,3	6	40	6	40
2. The accuracy of the line type and thickness used in the project	-	-	3	20	8	53,3	3	20	1	6,66	-	-	-	-	4	26,6	7	46,6	4	26,6
3. Compliance of drawing elements with the regulation places	-	-	2	13,3	11	73,3	2	13,3	-	-	-	-	1	6,6	-	-	5	33,3	9	60
4. Suitability of regulation with lamp location, number and forces	1	6,6	4	26,6	6	40	4	26,6	-	-	-	-	-	-	4	26,6	5	33,3	6	40
5. Conformity of the socket number and forces with regulation	-	-	2	13,3	5	33,3	6	40	2	13,3	-	-	-	-	3	20	8	53,3	4	26,6
6. Adequacy of management in the drawing layer	-	-	4	26,6	7	46,6	3	20	1	6,66	-	-	-	-	4	26,6	3	20	8	53,3
7.Accuracy of the strong current colon chart	-	-	5	33,3	3	20	6	40	1	6,66	-	-	2	13,3	6	40	3	20	4	26,6
8. Accuracy of the low current colon chart	-	-	4	26,6	5	33,3	6	40	-	-	-	-	1	6,6	5	33,3	4	26,6	5	33,3
9. Accuracy of calculation of the voltage drop	2	13,3	4	26,6	7	46,6	2	13,3	-	-	-	-	-	-	6	40	4	26,6	5	33,3
10. Conformity of scaling	-	-	1	6,6	7	46,6	5	33,3	2	13,3	-	-	-	-	4	26,6	8	53,3	3	20
11. Accuracy of the upload table	-	-	5	33,3	5	33,3	5	33,3	-	-	-	-	-	-	7	46,6	3	20	5	33,3
12. Accuracy of calculation of the grounding	-	-	4	26,6	6	40	3	20	2	13,3	-	-	-	-	3	20	5	33,3	7	46,6
13. The correct use of time in the project execution process	5	33,3	8	53,3	2	13,3	-	-	-	-	-	-	-	-	3	20	6	40	6	40

IV. CONCLUSION

In recent years, the effects from computer-assisted education on learning is extensively examined by researchers in different fields. Computer-assisted education is been used frequently in modern educational systems because of its benefits like providing persistence in learning in general, providing a learner-centered learning process, getting the event of learning out of four walls and making it independent from space and time, providing the possibility to practice frequently and providing a quick access to information.

In this experimental study, the effects from computer-assisted education on the success of students taking the course Electrical Installation Project drawing were researched. The results obtained in the light of research findings are presented in the following items:

According the results of the scale developed by researchers, was the academic success of the group that had received a visual training CD as a supplement to the traditional education was higher than the success of the group that had learned only with the traditional education system.

It is observed that the students of the group that possessed a private didactic software during a process of approximately one month, asked the instructors for less help. This shows that the visual education CD contributed to the development of the individual competences of the students.

According these results it can be defend that in project-based courses the targeted and specially developed visual and auditory didactic materials that the students can consult independent of time or place during the project are more effective than other materials in scattered places.

There may be some disadvantages besides the benefits that the study revealed. For example, the teamspirit may weaken because the students work individually. This and similar disadvantages are waiting for researchers as a subject of an other study.

REFERENCES

- [1] C. Alkan, H. Dogan & I. Sezgin, (1999) "Principles of Vocational and Technical Education", Gazi University Faculty of Communication Press, Ankara.
- [2] H. Ocal, (2008) "Vocational Education and Vocational Guidance", In the light of Science and Intellect Journal of Education, Vol. 99, pp. 12–19.
- [3] I. Esme, (2007) "Current Status and Issues about Vocational and Technical Education", T.C. YOK International Vocational and Technical Education Conference, Ankara, Turkey.

- [4] P. L. Smith & T. J. Ragan, (1999) "Instructional Design", Second Edition, USA: John Wiley&Sons Inc.
- [5] S. Cengizhan, (2007) "The Effects Of Project Based And Computer Assisted Instructional Designs On Those Students' Who Have Depended, Independed And Cooperative Learning Styles, Academic Achievement And Learning Retention" Vol. 5, No. 3, pp. 377–401.
- [6] N. Bajraktarevic, W. Hall & P. Fullick, (2003) "Incorporating Learning Styles in Hypermedia Environment: Empirical Evaluation", Accessed: <http://www.wis.win.tue.nl/ah2003/proceedings/paper4.pSd>, April-2005.
- [7] J. Ingham, R. Meza, P. Miriam & G. Price, (1998) "A Comparison of the Learning Style and Creative Talents of Mexican and American Undergraduate Engineering Students", Accessed: <http://fie.engr.pitt.edu/fie98/papers/1352.pSd>, January-2004.
- [8] R. M. Felder, (1996) "Matters of Style", ASEE Prism, Vol. 6, No. 4, pp. 18–23.
- [9] O. Demirbas & H. Demirkan, (2003) "Focus on Architectural Design Process Through Learning Style, Design Studies", Vol. 24, No 5, pp. 437–456.
- [10] R. Dunn & S. Griggs, (1996) "Hispanic-American Students and Learning Style", ERIC Identifier:ED393607, Accessed: <http://www.ericfacility.net/ericdigests/ed393607.html>, January-2004.
- [11] L. A. Kraus, W. M. Reed & G. E. Fitzgerald, (2001) "The Effects Of Learning Style and Hypermedia Prior Experience On Behavioral Disorders Knowledge And Time On Task: A Case-Based Hypermedia Environment, Computers in Human Behavior", Vol. 17, No. 1, pp. 124–140.
- [12] I. Sahin & T. Findik, (2008) "Vocational and Technical Education in Turkey: Current Situation, problems and proposition for solutions", The Turkish Journal of Social Research Vol. 12, No. 3, pp. 66–86.
- [13] H. Ogut, A. A. Altun, S. A. Sulak & H. E. Kocer, (2004) "Computer-assisted, Interne Access, E-Learning with Interactive Training Cd", The Turkish Online Journal of Educational Technology, Vol. 3, No 1, pp. 67–74.
- [14] B. Arslan, (2003) "Computer-assisted Education Secondary Students and the feedback from CAE of teachers which have had an educative role in this process", The Turkish Online Journal of Educational Technology, Vol. 2, No. 4, pp. 67–75.
- [15] F. Odabasi, (2006) "Computer-assisted education", Unit 8, Anadolu University, Open Education Faculty Press, Vol. 135, Eskisehir.
- [16] K. Cilenti, (1995) "Educational Technology and Importance", Kadioglu Printing, Ankara.
- [17] C. Alkan, D. Deryakulu & N. Simsek, (1995) "Introduction to Educational Technology: Discipline, Process, Product", Onder Printing, Ankara.
- [18] B. Akkoyunlu, (1998) "The place and role of the teacher in the Curriculum Programs of Computers", Hacettepe University Press, Ankara.
- [19] M. Buyukbayraktar, (2006) "The effect of computer assisted application of the Logical Circuit Design on the succes of students", Unpublished Master's Thesis, Sakarya University, Institute of Social Sciences, Sakarya.
- [20] H. Ekiz, Y. Bayam & H. Unal (2003) "Application of distance education on logical circuits", The Turkish Online Journal of Educational Technology, Vol. 2, No. 4, pp. 92–94.
- [21] A.H.K. Yuen (2006) "Learning to program through interactive simulation", Educational Media International, Vol. 43, No. 3, pp. 251–268.
- [22] H. Keser, (1988) "Proposed Model for Computer-assisted education", PhD Thesis, Ankara University, Institute of Social Sciences, Ankara.
- [23] G. Gleason, (1981) "Microcomputers in Education: The State of Art." Education Technology, Vol. 21, No. 3.
- [24] S. Usun, (2003) "The Views Of The Students On The Advantages And Important Using Elements In Education And Instruction Of Computers", Kastamonu University, Journal of Education, Vol. 1, No. 2, pp. 367–378.
- [25] E. Bayraktar, 1998 "Computer-assisted Teaching of Mathematics", PhD Thesis Ankara University, Institute of Social Sciences, Ankara.

Authors

İsmail Kayri is both a lecturer in Department of Electric Education in Batman University and a PhD scholar in Electric-Electronic Department in Firat University. He studies on electric systems installation, programming language, database management system and other software tools especially related to electric & electronic science.



Dr. Muhsin Tunay Gencoglu is an associate professor in Department of Electrical and Electronics Engineering, Firat University. Fields of interest: high voltage techniques, electric power transmission and disturbance, HV insulators, lighting techniques and renewable energy sources. He has lots of articles in these fields.



Dr. Murat Kayri is an assistant professor in Computer Science and Instructional Technology Department in Yuzuncu Yil University. Dr. Kayri interests in neural network, statistical modelling, and networking. He has lots of articles on statistical and artificial neural network.



FRACTAL CHARACTERIZATION OF EVOLVING TRAJECTORIES OF DUFFING OSCILLATOR

Salau, T. A.O.¹ and Ajide, O.O.²

^{1,2} Department of Mechanical Engineering, University of Ibadan, Nigeria.

ABSTRACT

This study utilised fractal disk dimension characterization to investigate the time evolution of the Poincare sections of a harmonically excited Duffing oscillator. Multiple trajectories of the Duffing oscillator were solved simultaneously using Runge-Kutta constant step algorithms from set of randomly selected very close initial conditions for three different cases. These initial conditions were from a very small phase space that approximates geometrically a line. The attractor highest estimated fractal disk dimension was first recorded at the end of 15, 22, and 5 excitation periods for Case-1, Case-2 and Case-3 respectively. The corresponding scatter phase plots for Case-1 and Case-2 agreed qualitatively with stroboscopic-ally obtained Poincare sections found in the literature. The study thus established sensitivity of Duffing to initial conditions when driven by different combination of damping coefficient, excitation amplitude and frequency. It however showed a faster, accurate and reliable alternative computational method for generating its Poincare sections.

KEYWORDS: *Duffing oscillator, Fractal, Poincare sections, Trajectories, Disk dimension, Runge-Kutta and phase space*

I. INTRODUCTION

Duffing oscillator can be described as an example of a periodically forced oscillator with a nonlinear elasticity [14]. This can be considered as chaotic system since it is characterized by nonlinearity and sensitivity to initial conditions. Available literature shows that Duffing oscillator has been highly studied and this is due to its wide modelling applications in various fields of dynamics. The dynamics of duffing oscillator has been studied using various tools. [9] investigated the dynamical behaviour of a duffing oscillator using bifurcation diagrams. The results of the study revealed that while bifurcation diagram is a resourceful instrument for global view of the dynamics of duffing oscillator system over a range of control parameter, it also shows that its dynamics depend strongly on initial conditions. [11] Investigated the dynamic stabilization in the double-well Duffing oscillator using bifurcation diagrams. The research paper identified an interesting behaviour in the dynamic stabilization of the saddle fixed point. It was observed that when the driving amplitude is increased through a threshold value, the saddled fixed point. It was observed that when the driving amplitude is increased through a threshold value, the saddle fixed point becomes stabilized with the aid of a pitchfork bifurcation. The findings of the authors revealed that after the dynamic stabilization, the double-well Duffing oscillator behaves as the single –well Duffing oscillator. This is because the effect of the central potential barrier on the dynamics of the system becomes negligible.

A fractal generally refers to a rough or fragmented geometric shape which is capable of been divided into parts. Each part is an approximately reduced-size copy of the whole. This property is popularly referred to as ‘self-similarity’. We can also describe fractal as geometric pattern that is repeated at ever smaller scales to produce irregular shapes and surfaces that cannot be represented by classical geometry. The complex nature of fractal is becoming to attract more researchers interest in the recent time. This is because it has become a major fundamental of nonlinear dynamics and theory of chaos.

Fractal structures and dynamical systems associated with phase plots are inseparable. The strong relationship between fractal structures and chaos theory will continue to remain the platform of success in nonlinear dynamics. Fractals are highly employed in computer modelling of irregular patterns and structures in nature. Though the theory of chaos and the concept of fractals evolved independently, they have been found to penetrate each other's front. The orbits of nonlinear dynamical system could be attracted or repelled to simple shape of nonlinear, near-circles or other shapes of Elucid[10]. He furthered his explanation that, however, these are rare exceptions and the behaviour of most nonlinear dynamical systems tends to be more complicated. The analysis of nonlinear dynamics fractal is useful for obtaining information about the future behaviour of complex systems [5]. The main reason for this is because they provide fundamental knowledge about the relation between these systems and uncertainty and indeterminism. [5] research paper focus on fractal structures in nonlinear dynamics. The work clearly describes the main types of fractal basin, their nature and the numerical and experimental techniques used to obtain them from both mathematical models and real phenomena. [7] Research paper was on intermingled fractal Arnold tongues. The paper presented a pattern of multiply interwoven Arnold tongues in the case of the single-well Duffing oscillator at low dissipation and weak forcing. It was observed that strips $2/2$ Arnold tongues formed a truncated fractal and the tongue-like regions in between are filled by finely intermingled fractal like $1/1$ and $3/3$ Arnold tongues, which are fat fractals characterized by the uncertainty exponent α approximate to 0.7. The findings of authors showed that the truncated fractal Arnold tongues is present in the case of high dissipation as well, while the intermingled fractal pattern gradually disappears with increasing dissipation. [16] Research paper was on $1/3$ pure sub-harmonic solution and fractal characteristic of transient process for Duffing's equation. The investigation was carried out using the methods of harmonic balance and numerical integration. The author introduced assumed solution and was able to find the domain of sub-harmonic frequencies. The asymptotical stability of the sub-harmonic resonances and the sensitivity of the amplitude responses to the variation of damping coefficient were examined. Then, the subatomic resonances were analyzed by using techniques from the general fractal theory. The analysis reveals that the sensitive dimensions of the system time-field responses show sensitivity to the conditions of changed initial perturbation, changed damping coefficient or the amplitude of excitation. The author concluded that the sensitive dimension can clearly describe the characteristics of the transient process of the subharmonic resonances. According to [15], the studies of the phenomenon of chaos synchronization are usually based upon the analysis of transversely stable invariant manifold that contains an invariant set of trajectories corresponding to synchronous motions. The authors developed a new approach that relies on the notions of topological synchronization and the dimension for Poincare recurrences. The paper showed that the dimension of Poincare recurrences may serve as an indicator for the onset of synchronized chaotic oscillations. The hallmark of [12] paper in 2007 was to examine the application of a simple feedback controller to eliminate the chaotic behaviour in a controlled extended Duffing system. The reason was to regulate the chaotic motion of an extended Duffing system around less complex attractors, such as equilibrium points and periodic orbits. The author proposed a feedback controller which consists of a high-pass filter and a saturator. This gives the opportunity of simple implementation and can be made on the basis of measured signals. The authors sufficiently demonstrated this feedback control strategy using numerical simulations. [8] Study was on characterization of non stationary chaotic systems. The authors noticed that significant work has not been done in the characterization of these systems. The paper stated that the natural way to characterize these systems is to generate and examine ensemble snapshots using a large number of trajectories, which are capable of revealing the underlying fractal properties of the system. The authors concluded that by defining the Lyapunov exponent and the fractal dimension based on a proper probability measure from the ensemble snapshots, the Kaplan-Yorke formula which is fundamental in nonlinear dynamics can be shown. This finding remains correct most of the time even for non-stationary dynamical systems.

Chaotic dynamical systems with phase space symmetries have been considered to exhibit riddle basins of attraction [1]. This can be viewed as extreme fractal structures not minding how infinitesimal the uncertainty in the determination of an initial condition. The authors noticed that it is not possible to decrease the fraction of such points that will surely converge to a given attractor. The main aim of

the authors' work was to investigate extreme fractal structures in chaotic mechanical systems. The authors investigated mechanical systems depicting riddle basins of attraction. That is, a particle under two-dimensional potential with friction and time-periodic forcing. The authors were able to verify this riddling by checking its mathematical requirements through computation of finite-time Lyapunov exponents as well as by scaling laws that explain the fine structure of basin filaments densely intertwined in phase space. A critical characterization of non-ideal oscillators in parameter space was carried out by [13]. The authors investigated dynamical systems with non-ideal energy source. The chaotic dynamics of an impact oscillator and a Duffing oscillator with limited power supply were analyzed in two-dimensional parameter space by using the largest Lyapunov exponents identifying self-similar periodic sets, such as Arnold tongues and shrimp-like structures. For the impact oscillator, the authors identified several coexistence of attractors showing a couple of them, with fractal basin boundaries. According to the paper, these kinds of basin structures introduce a certain degree of unpredictability on the final state. The simple interpretation of this is that the fractal basin boundary results in a severe obstruction to determine what attractor will be a fine state for a given initial condition with experimental error interval.

Fractal characterization of evolving trajectories of a dynamical system will no doubt be of immense help in diagnosing the dynamics of very important chaotic systems such as Duffing oscillator. Extensive literature search shows that disk dimension is yet to be significantly employed in fractal characterization of Duffing oscillator. The objective of this study is to investigate and characterize the time evolution of Poincare sections of a harmonically excited Duffing oscillator using fractal disk dimension.

This article is divided into four sections. Section 1 gives the study background and brief review of literature. Section 2 gives the detail of methodology employed in this research. Subsection 2.1 gives the equation of harmonically excited duffing oscillators that is employed in demonstrating fractal characterization of evolving trajectories. Subsection 2.1 gives explanation on the parameter details of all the studied cases. Different combinations of damping coefficient and excitation amplitude considered are clearly stated. The methodology is concluded in subsection 2.3 where explanation is given on how attractor is characterized. Section 3 gives detail results and discussion. The findings of this work are summarized in section 4 with relevant conclusions.

II. METHODOLOGY

2.1 Duffing Oscillator

The studied normalized governing equation for the dynamic behaviour of harmonically excited Duffing system is given by equation (1).

$$\ddot{x} + \gamma \dot{x} - \frac{x}{2}(1 - x^2) = P_o \sin(\omega t) \quad (1)$$

In equation (1) x , \dot{x} and \ddot{x} represents respectively displacement, velocity and acceleration of the Duffing oscillator about a set datum. The damping coefficient is γ . Amplitude strength of harmonic excitation, excitation frequency and time are respectively P_o , ω and t . [2], [3] and [6] proposed that combination of $\gamma = 0.168$, $P_o = 0.21$, and $\omega = 1.0$ or $\gamma = 0.0168$, $P_o = 0.09$ and $\omega = 1.0$ parameters leads to chaotic behaviour of harmonically excited Duffing oscillator. This study investigated the evolution of 3000 trajectories that started very close to each other and over 25 excitation periods at a constant step ($\Delta t = \frac{\text{Excitation period}}{500}$) in Runge-Kutta fourth order algorithms. The resulting

attractors (see [4]) at the end of each excitation period were characterized with fractal disk dimension estimate based on optimum disk count algorithms.

2.2 Parameter details of studied cases

Three different cases were studied using the details given in table 1 in conjunction with governing equation (1). Common parameters to all cases includes initial displacement range ($0.9 \leq x \leq 1.1$),

Zero initial velocity (\dot{x}), excitation frequency (ω) and random number generating seed value of 9876.

Table 1: Combined Parameters for Cases

Cases	Damping coefficient (γ)	Excitation amplitude (P_o)
Case-1	0.1680	0.21
Case-2	0.0168	0.09
Case-3	0.0168	0.21

2.3 Attractor Characterization

The optimum disk count algorithm was used to characterize all the resulting attractors based on fifteen (15) different disk scales of examination and over five (5) independent trials.

III. RESULTS AND DISCUSSION

The scatter phase plots of figures 1, 2 and 3 shows the comparative attractors resulting from the time evolution of trajectories of Duffing oscillator for the studied cases.

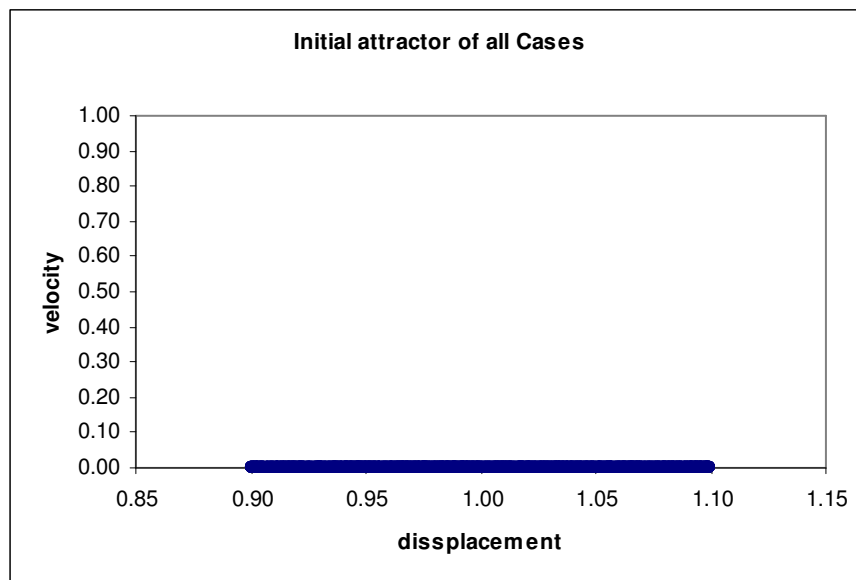


Figure 1: Attractor of all cases at zero excitation period.

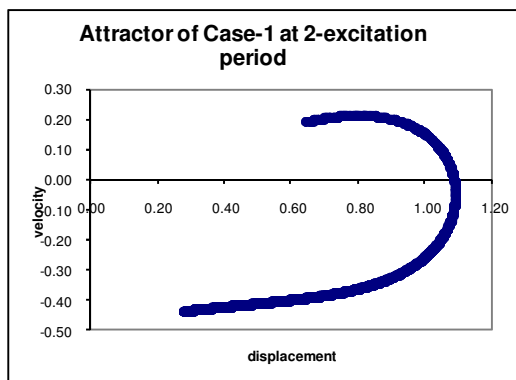


Fig. 2 (a)

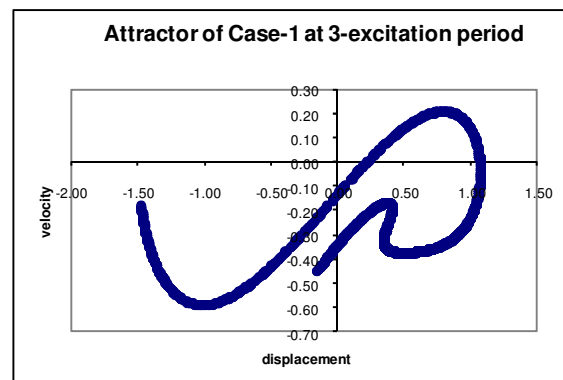


Fig. 2 (b)

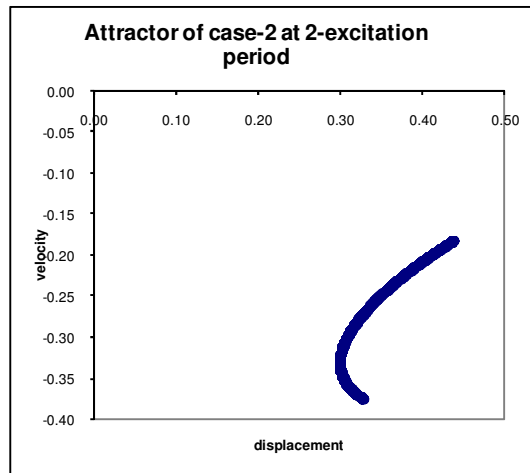


Fig. 2 (c)

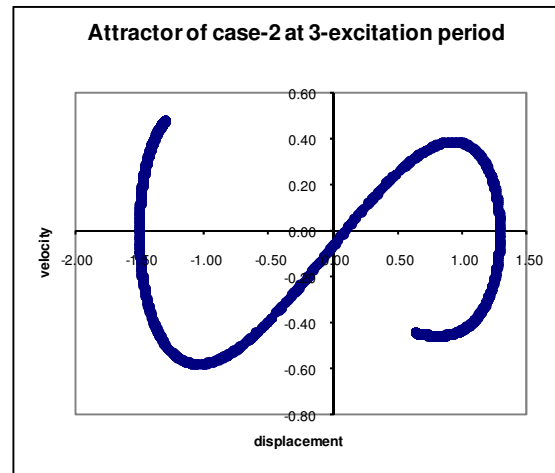


Fig. 2 (d)

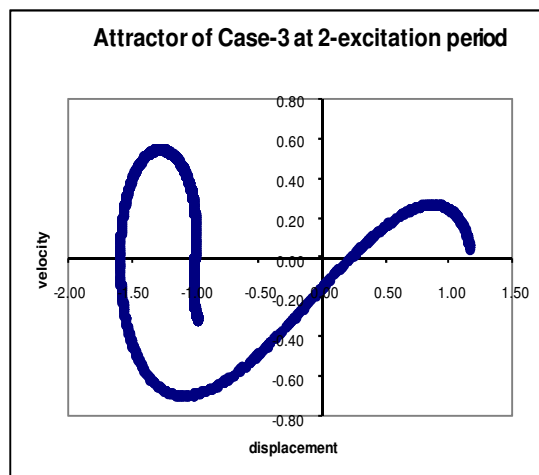


Fig. 2 (e)

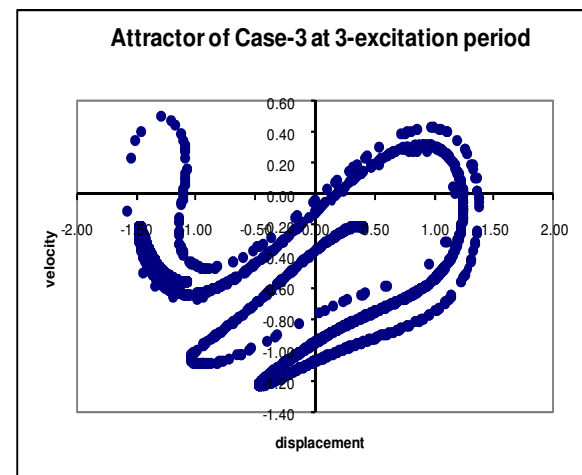


Fig. 2 (f)

Figure 2: Comparison of attractors at 2 and 3 excitation periods.

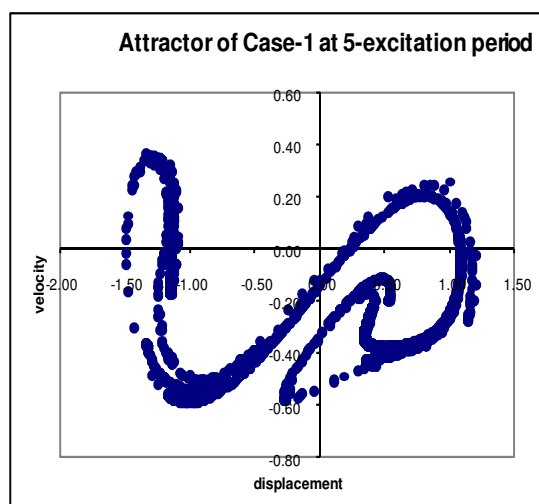


Fig. 3 (a)

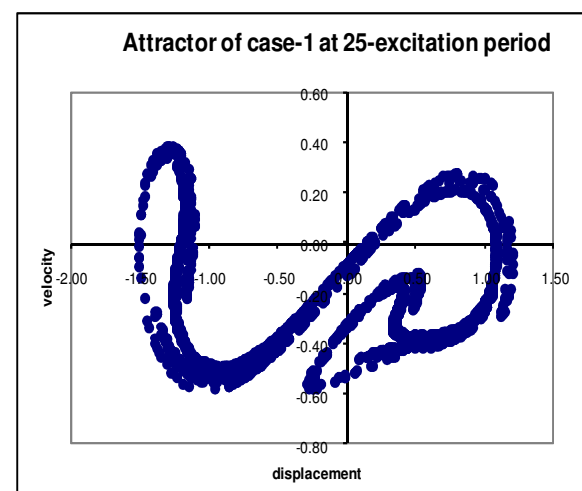


Fig. 3 (b)

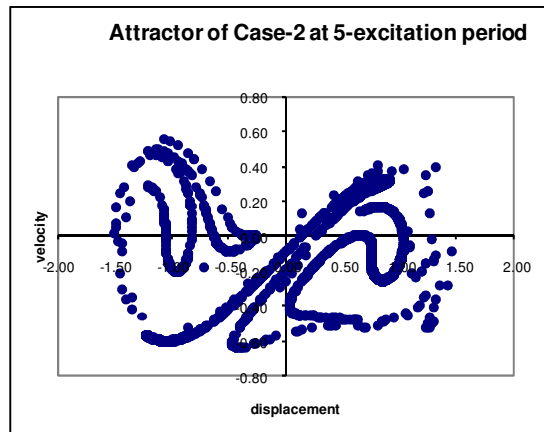


Fig. 3 (c)

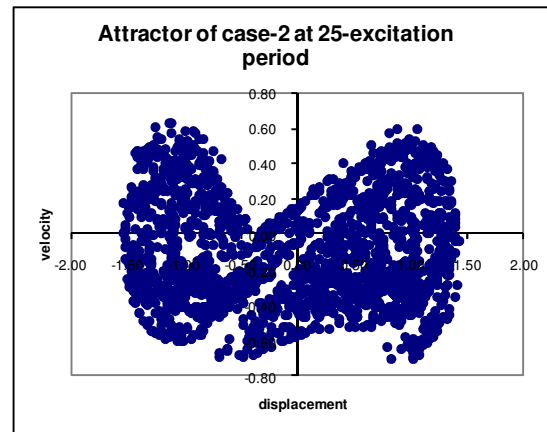


Fig. 3 (d)

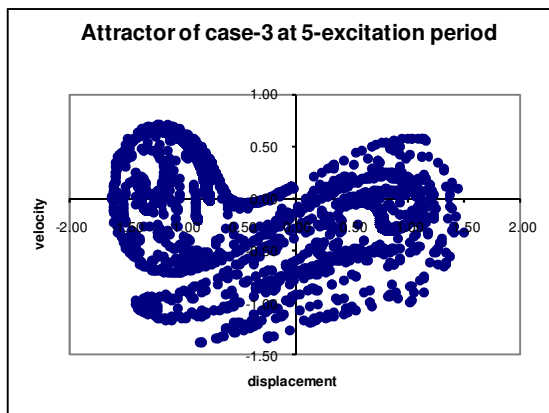


Fig. 3 (e)

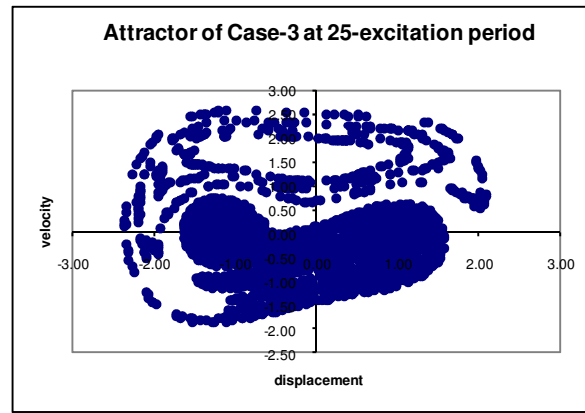


Fig. 3 (f)

Figure 3: Comparison of attractors at 5 and 25 excitation periods.

Referring to figures 1, 2 and 3 the geometrical complexity of the attractors varied widely with cases and number of excitation periods. This is an affirmation of high sensitivity to initial conditions of Duffing oscillator behaviour if excited harmonically by some parameters combinations. The attractors of Case-1 and Case-2 approach qualitatively their respective stroboscopic-ally obtained Poincare section with increasing excitation period.

The varied geometrical complexity of the attractors presented in figures 1, 2, and 3 can be characterized using fractal disk dimension measure. The algorithms for estimating the fractal disk dimension is demonstrated through presentation in table 2 and figure 4.

Table 2: Disk required for complete cover of Case-1 attractor (Poincare section) at the end of 25 excitation periods.

Disk scale	Optimum Disk counted	Disk counted in five (5) trials				
		1	2	3	4	5
1	2	3	2	2	2	2
2	4	5	4	4	5	4
3	6	8	6	8	8	8
4	11	14	12	12	12	11
5	17	19	18	18	17	17
6	21	21	21	22	21	21
7	25	25	28	27	26	28
8	28	31	31	28	30	31
9	34	38	37	34	39	37
10	40	40	42	45	41	43

11	45	47	47	49	46	45
12	52	54	53	55	52	54
13	60	60	62	61	60	62
14	61	65	65	67	64	61
15	68	72	69	69	72	68

Table 2 refers physical disk size for disk scale number one (1) is the largest while disk scale number fifteen (15) is the smallest. The first appearances of the optimum disk counted in five independent trials are shown in bold face through the fifteen scales of examination. Thus the optimum disk counted increases with decreasing physical disk size. The slope of line of best fit to logarithm plots of corresponding disk scale number and optimum disk counted gives the estimated fractal disk dimension of the attractor. Referring to figure 4 the estimated fractal dimension of the attractor of Case-1 at the end of 25-excitation periods is 1.3657 with an R^2 value of 0.9928.

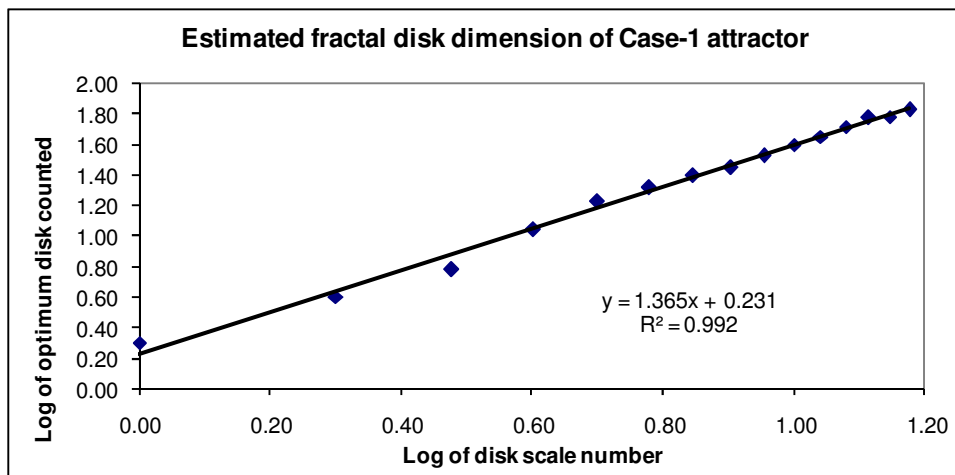


Figure 4: Fractal disk dimension of case-1 attractor at the end of 25 excitation periods.

The variation of estimated fractal disk dimension of attractors for studied cases with increasing excitation period is given in figure 5.

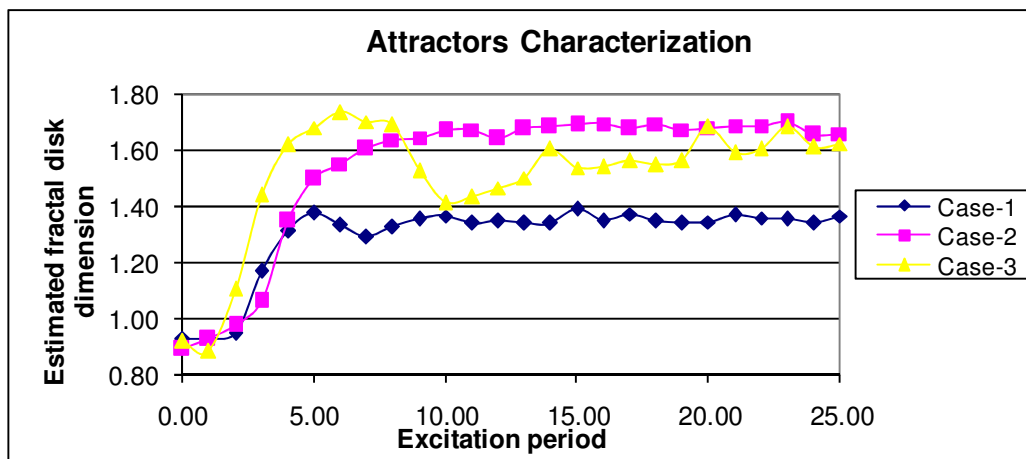


Figure 5: Variation of estimated disk dimension of attractors with excitation period.

Figure 5 refers a rise to average steady value of estimated fractal disk dimension was observed for all studied cases except Case-3. This observation with Case-3 may be due to its low damping value ($\gamma=0.0168$) and relative very high excitation amplitude ($P_o=0.21$). The attractor highest estimated fractal disk dimension of 1.393, 1.701 and 1.737 was recorded for the first time at corresponding excitation periods of 15, 23 and 5 for Case-1, Case-2 and Case-3 respectively.

Table 3: Estimated fractal disk dimension of Case-1 attractors at the end of 26-different excitation periods.

Standard deviation	Excitation period	Case-1 attractor different estimated fractal disk dimensions						
		Optimum	Average	Five different trials				
				1	2	3	4	5
0.02	0	0.928	0.903	0.898	0.896	0.878	0.924	0.919
0.02	1	0.927	0.918	0.917	0.908	0.920	0.888	0.956
0.01	2	0.948	0.938	0.929	0.949	0.956	0.933	0.923
0.03	3	1.170	1.146	1.128	1.161	1.137	1.121	1.182
0.03	4	1.314	1.259	1.262	1.285	1.205	1.261	1.284
0.03	5	1.376	1.340	1.351	1.348	1.308	1.315	1.380
0.02	6	1.333	1.305	1.275	1.315	1.317	1.293	1.327
0.01	7	1.292	1.297	1.307	1.304	1.292	1.293	1.290
0.01	8	1.325	1.327	1.331	1.344	1.312	1.323	1.328
0.02	9	1.355	1.341	1.358	1.309	1.351	1.332	1.357
0.02	10	1.368	1.333	1.319	1.377	1.331	1.323	1.317
0.02	11	1.341	1.324	1.323	1.350	1.348	1.295	1.306
0.02	12	1.350	1.335	1.309	1.326	1.369	1.349	1.320
0.02	13	1.344	1.341	1.330	1.357	1.361	1.348	1.310
0.02	14	1.339	1.314	1.330	1.296	1.282	1.333	1.328
0.03	15	1.394	1.345	1.324	1.324	1.325	1.400	1.351
0.02	16	1.350	1.332	1.309	1.324	1.348	1.361	1.320
0.02	17	1.374	1.345	1.361	1.356	1.362	1.327	1.320
0.03	18	1.349	1.332	1.313	1.356	1.371	1.332	1.290
0.01	19	1.343	1.341	1.325	1.357	1.341	1.333	1.352
0.06	20	1.346	1.319	1.335	1.357	1.356	1.216	1.331
0.04	21	1.368	1.340	1.344	1.355	1.270	1.341	1.390
0.02	22	1.359	1.342	1.355	1.318	1.319	1.344	1.375
0.05	23	1.356	1.323	1.342	1.331	1.335	1.362	1.242
0.02	24	1.342	1.331	1.329	1.358	1.305	1.315	1.345
0.06	25	1.366	1.331	1.229	1.383	1.361	1.325	1.356

Table 4: Estimated fractal disk dimension of Case-2 attractors at the end of 26-different excitation periods.

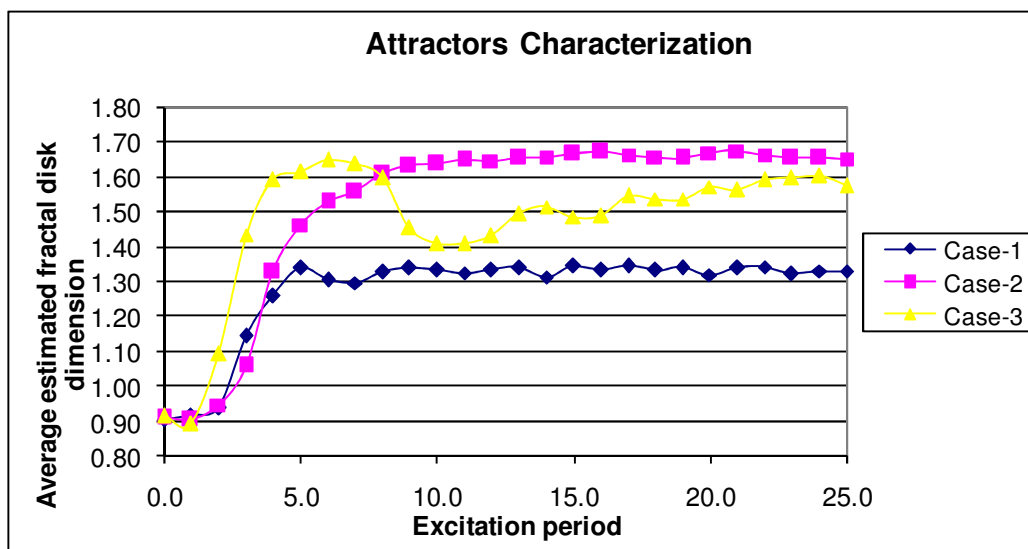
Standard deviation	Excitation period	Case-2 attractor different estimated fractal disk dimensions						
		Optimum	Average	Five different trials				
				1	2	3	4	5
0.01	0	0.889	0.910	0.924	0.909	0.898	0.896	0.923
0.03	1	0.926	0.906	0.920	0.890	0.894	0.883	0.944
0.06	2	0.975	0.948	0.955	0.984	0.845	0.976	0.979
0.01	3	1.063	1.058	1.063	1.059	1.057	1.041	1.071
0.02	4	1.347	1.326	1.330	1.334	1.308	1.353	1.308
0.02	5	1.499	1.463	1.481	1.495	1.452	1.449	1.437
0.02	6	1.552	1.528	1.513	1.549	1.515	1.540	1.520
0.05	7	1.605	1.558	1.567	1.621	1.554	1.480	1.571
0.01	8	1.638	1.609	1.596	1.605	1.598	1.617	1.626
0.02	9	1.646	1.630	1.626	1.653	1.601	1.643	1.629
0.02	10	1.669	1.636	1.616	1.666	1.622	1.627	1.647
0.02	11	1.674	1.648	1.667	1.650	1.621	1.651	1.650
0.01	12	1.644	1.646	1.630	1.657	1.642	1.642	1.656
0.03	13	1.678	1.653	1.669	1.690	1.631	1.637	1.639
0.02	14	1.683	1.658	1.671	1.658	1.658	1.626	1.676
0.02	15	1.691	1.664	1.669	1.650	1.702	1.661	1.639
0.01	16	1.697	1.671	1.665	1.685	1.659	1.670	1.673
0.01	17	1.679	1.664	1.675	1.653	1.652	1.653	1.683
0.05	18	1.696	1.657	1.695	1.682	1.577	1.680	1.654
0.01	19	1.675	1.655	1.653	1.657	1.655	1.641	1.667
0.02	20	1.682	1.669	1.676	1.659	1.635	1.683	1.694
0.02	21	1.688	1.675	1.681	1.707	1.674	1.667	1.648

0.02	22	1.688	1.664	1.637	1.664	1.664	1.661	1.695
0.05	23	1.701	1.656	1.712	1.583	1.643	1.681	1.661
0.01	24	1.656	1.656	1.665	1.660	1.664	1.630	1.661
0.01	25	1.660	1.652	1.665	1.660	1.647	1.650	1.636

Table 5: Estimated fractal disk dimension of Case-3 attractors at the end of 26-different excitation periods.

Standard Deviation	Excitation period	Case-3 attractor different estimated fractal disk dimensions						
		Optimum	Average	Five different trials				
				1	2	3	4	5
0.03	0	0.917	0.915	0.895	0.905	0.889	0.945	0.943
0.01	1	0.881	0.892	0.895	0.898	0.910	0.883	0.875
0.02	2	1.107	1.094	1.122	1.079	1.108	1.090	1.073
0.02	3	1.441	1.434	1.457	1.430	1.436	1.415	1.434
0.02	4	1.619	1.593	1.597	1.620	1.584	1.583	1.584
0.05	5	1.682	1.615	1.592	1.654	1.683	1.589	1.558
0.04	6	1.737	1.648	1.623	1.635	1.628	1.724	1.631
0.05	7	1.704	1.636	1.564	1.671	1.651	1.619	1.675
0.04	8	1.695	1.598	1.536	1.610	1.601	1.632	1.609
0.04	9	1.527	1.453	1.434	1.433	1.429	1.530	1.440
0.02	10	1.415	1.408	1.380	1.411	1.423	1.419	1.409
0.03	11	1.432	1.410	1.361	1.421	1.412	1.422	1.434
0.04	12	1.467	1.432	1.470	1.409	1.385	1.440	1.458
0.02	13	1.504	1.495	1.506	1.508	1.497	1.494	1.469
0.04	14	1.605	1.514	1.510	1.501	1.505	1.576	1.478
0.04	15	1.540	1.486	1.495	1.457	1.437	1.543	1.496
0.05	16	1.541	1.490	1.465	1.445	1.461	1.541	1.536
0.02	17	1.562	1.545	1.552	1.543	1.508	1.554	1.566
0.01	18	1.551	1.538	1.548	1.556	1.528	1.530	1.529
0.03	19	1.565	1.536	1.489	1.536	1.543	1.548	1.566
0.04	20	1.683	1.571	1.634	1.545	1.565	1.565	1.545
0.02	21	1.592	1.561	1.574	1.564	1.564	1.528	1.575
0.02	22	1.606	1.590	1.617	1.577	1.606	1.569	1.581
0.06	23	1.687	1.599	1.586	1.603	1.576	1.695	1.534
0.01	24	1.614	1.603	1.584	1.618	1.610	1.599	1.603
0.03	25	1.623	1.576	1.607	1.556	1.606	1.525	1.584

Tables 3, 4 and 5 refers the variation of optimum estimated fractal disk dimension with increasing excitation period is shown in figure 5.

**Figure 6:** Variation of average estimated fractal disk dimension of attractors with excitation period.

In addition the variation of average estimated fractal disk dimension based on five independent trials with increasing excitation period is shown in figure 6. Figures 5 and 6 are same qualitatively. However the average estimated fractal disk dimensions are consistently lower than the corresponding optimum estimated fractal disk dimension for all attractors characterized. Standard deviation estimated for five trial results lies between minimum of 0.01 and maximum of 0.06 for all the cases and the attractors.

Figures 5 and 6 indicated that the attractors for different cases ultimately evolve gradually to steady geometric structure.

IV. CONCLUSIONS

The study has demonstrated the Duffing oscillator high sensitivity behaviour to set of very close initial conditions under the combination of some harmonic excitation parameters. Cases 1 and 2 evolve gradually to unique attractors which are comparable to corresponding Poincare sections obtained in the literature. On the final note, this study establishes the utility of fractal dimension as effective characterization tool and a novel alternative computational method that is faster, accurate and reliable for generating Duffing attractors or Poincare sections.

REFERENCES

- [1]. Carmago, S.; Lopes, S.R. and Viana. 2010. Extreme fractal structures in chaotic mechanical systems : Riddled basins attractor. XI Latin American workshop on nonlinear phenomena. Journal of physics: Conference series 246 (2010) 012001. Doi: 10.1088/1742-6596/246/1/012001. IOP publishing Ltd.
- [2]. Dowell, E.H. 1988. Chaotic oscillations in mechanical systems, Computational Mechanics, 3, 199-216.
- [3]. Francis, C.M. 1987. Chaotic Vibrations-An Introduction for Applied Scientists and Engineers, John Wiley & Sons, New York, ISBN 0-471-85685-1
- [4]. Gregory, L.B and Jerry, P.G. 1990. Chaotic dynamics: An introduction. Cambridge university Press, New York, ISBN 0-521-38258-0 Hardback, ISBN 0-521-38897-X Paperback.
- [5]. Jacobo, A.; Ricardo, L.V. and Miguel, A.F.S. 2009. Fractal structures in nonlinear dynamics. American physical society's new journal. Vol. 81, Issue 1.
- [6]. Narayanan, S. and Jayaraman, K. 1991. Chaotic vibration in a non-linear oscillator with coulomb damping. Journal of sound and vibration. Vol. 146, Issue 1, pg. 17-31. Published by Elsevier Ltd.
- [7]. Paar, V. and Pavin, N. 1998. Intermingle fractal Arnold tongues. Physical review. A. 57 (2 part A) : 1544-1549. ISSN 1050-2947.
- [8]. Ruth, S.; Ying-Cheng, L. and Qingfei, C. 2008. Characterisation of non stationary chaotic systems. American physical society new journal. Physical review E. Vol. 77, Issue 2.
- [9]. Salau, T.A.O. and Ajide, O.O. 2011. Investigating Duffing oscillator using bifurcation diagrams. International journal of mechanics structural. ISSN 097312X, Vol. 2, No. 2. pp. 57-68. house. <http://www.irphouse.com>.
- [10]. Salau, T.A.O. 2002. Fractal analyses and characterization of tread patterns of automobile tyres. Unpublished Doctor of philosophy (Ph.D) Thesis. Mechanical engineering department, University of Ibadan, Nigeria.
- [11]. Sang-Yoon, K. and Youngtae, K. 2000. Dynamic stabilization in the double-well Duffing oscillator. American physical society new's journal. Physical review E. Vol. 61, No. 6
- [12]. Sihem, A.L.; Samuel, B.; Kakmeni, F. M. M.; Brahim, C. and Nouredine Ghouali. 2007. Chaos control using small-amplitude damping signals of the extended Duffing equation. Communications in nonlinear science and numerical simulation. Vol. 12, Issue 5, pg. 804-813. Copyright: Elsevier B.V.
- [13]. Silvio, L.T.; Ibere, L.C.; Jose, M.B. and Reyolando, M.L.R.F. 2010. Characterisation of non-ideal oscillators in parameter space. DINCON'10. 9th Brazilian conference on dynamics control and their applications. Serra Negra, Sp- ISSN 2178-3667. Pg-95.
- [14]. Takashi, K. 2008. Duffing oscillator dynamics. Source: Scholarpedia, 3(3):6327. Takashi is a lecturer at the Mechanical engineering department, Kogakuin University, Japan.
- [15]. Valentine, S.A.; Wen-Wei, L. and Nikolai, F.R. 2000. Fractal dimension for Poincare recurrences as an indicator of synchronized chaotic regimes. International journal of bifurcation and chaos. Vol. 10, No. 10. 2323-2337. World scientific publishing company.
- [16]. Yu-Xi, X.; Hai-yan, H. and Bang-Chun, W. 2000. 1/3 Sub-harmonic solution and fractal characteristic of transient process for Duffing's equation. Applied mathematics and mechanics, vol. 27, No. 9, Springer. 1171-1176, DOI: 10. 1007/s 10483-006-0903-1.

AUTHORS BIOGRAPHY

SALAU Tajudeen Abiola Ogunniyi is a senior Lecturer in the department of Mechanical of Engineering, University of Ibadan, Nigeria. He joined the services of the University of Ibadan in February 1993 as Lecturer II in the department of Mechanical Engineering. By virtue of hard work, he was promoted to Lecturer 1 in 2002 and senior Lecturer in 2008. He had served the department in various capacities. He was the coordinator of the department for 2004/2005 and 2005/2006 Academic sessions. He was the recipient of M.K.O Abiola postgraduate scholarship in 1993/1994 academic session while on his Ph.D research programme in the University of Ibadan. Salau has many publications in learned journals and international conference proceedings especially in the area of nonlinear dynamics. He had served as external examiner in departments of Mechanical Engineering of some institutions of higher learning in the country and a reviewer/rapporteur in some reputable international conference proceedings. His area of specialization is solid mechanics with bias in nonlinear dynamics and chaos. Salau is a corporate member, Nigerian Society of Engineers (NSE). He is a registered Engineer by the council for Regulations of engineering in Nigeria.(COREN). He is happily married and blessed with children.



AJIDE Olusegun Olufemi is currently a Lecturer II in the department of Mechanical Engineering, University of Ibadan, Nigeria. He joined the services of the University of Ibadan on 1st December 2010 as Lecturer II. He had worked as the Project Site Engineer/Manager of PRETOX Engineering Nigeria Ltd, Nigeria. Ajide obtained B.Sc (Hons.) in 2003 from the Obafemi Awolowo University, Nigeria and M.Sc in 2008 from the University of Ibadan, Nigeria. He received the prestigious Professor Bamiro Prize (Vice Chancellor Award) in 2008 for the overall best M.Sc student in Mechanical Engineering (Solid Mechanics), University of Ibadan, Nigeria. He has some publications in learned journals and conference proceedings. His research interests are in area of Solid Mechanics, Applied Mechanics and Materials Engineering. Ajide is a COREN registered Engineer. He is a corporate member of the Nigerian Society of Engineers (NSE) as well as corporate member of the Nigerian Institution of Mechanical Engineers (NIMechE).



SANKEERNA: A LINEAR TIME, SYNTHESIS AND ROUTING AWARE, CONSTRUCTIVE VLSI PLACER TO ACHIEVE SYNERGISTIC DESIGN FLOW

Santeppa Kambham¹ and Siva Rama Krishna Prasad Kolli²

¹ANURAG, DRDO, Kanchanbagh, Hyderabad, India

²ECE Dept, National Institute of Technology, Warangal, India

ABSTRACT

Standard cell placement is a NP complete open problem. The main objectives of a placement algorithm are to minimize chip area and the total wire length of all the nets. Due to interconnect dominance, Deep Sub Micron VLSI design flow does not converge leading to iterations between synthesis and layout steps. We present a new heuristic placement algorithm called Sankeerna, which tightly couples synthesis and routing and produces compact routable designs with minimum area and delay. We tested Sankeerna on several benchmarks using 0.13 micron, 8 metal layer, standard cell technology library. There is an average improvement of 46.2% in delay, 8.8% in area and 114.4% in wire length when compared to existing placement algorithms. In this paper, we described the design and implementation of Sankeerna algorithm and its performance is illustrated through a worked out example.

KEYWORDS: Placement, VLSI Design flow, Synthesis, Routing, Area and delay minimization

I. INTRODUCTION

VLSI chip complexity has been increasing as per the Moore's law, demanding more functionality, high performance, but with less design time. Producing compact layouts with high performance in shorter time is required, in order to meet the time to market needs of today's VLSI chips. This calls for tools, which run faster and also which converge without leading to design iterations. Placement is the major step in VLSI Design flow which decides the area and performance of the circuit. Detailed Routing is another time consuming step, which is performed after placement. If placement is not wire planned, routing may lead to heavy congestion resulting in several Design Rule Check (DRC) violations. It is required to iterate again with a new placement. If the wiring is not planned properly during placement, circuits routed may not meet the timing goals of the design. So there is a need for placers which are faster, produce compact layouts, meet the timing requirements and make the routing converge without DRC violations. The back end process of VLSI Design flow, that is, preparation of layout, is also to be tightly coupled with the front end synthesis process to avoid design iterations between synthesis and layout steps. It has been found that even after several iterations, this two step process does not converge and using wire load models this timing closure problem [1, 2] will not be solved.

In general, the standard cell placement problem can be stated as: Given a circuit consisting of technology mapped cells with fixed height and variable width, and a netlist connecting these cells, and Primary Inputs and Primary Outputs, construct a layout fixing the position of each cell without overlap with each other. The placement when routed should have minimum area, wire length, delay and should be routable. Minimum area is the area close to the sum of mapped standard cell areas.

Minimum wire length is the sum of all nets in the circuit when placed and routed. Delay is the delay of worst path in the routed circuit. Routability indicates that the layout should not be congested; wires routed should respect the Design Rules of the particular technology such that the routing is completed. Standard cell placement is known to be a NP complete open problem [3].

A Synergistic approach towards Deep Sub Micron (DSM) design, coupling logic synthesis and physical design is the need of the day [4, 1, 5]. There have been efforts to integrate synthesis and layout steps [6, 7, 8, 9]. All these efforts try to estimate wire delays, with the hope that they will be met finally, which is not happening. Wire delays are inevitable. The problem is not with wire delays, but with the non convergence and unpredictability. What we need is a quick way of knowing the final delay and a converging design flow. We have developed a design flow and a placer called *Sankeerna* targeted to produce compact routable layouts without using wire load models.

In Section 2, we briefly review the existing methods of placement and their limitations with respect to achieving a tightly coupled convergent design flow. Section 3 gives the basis for the *Sankeerna* algorithms. With this background, a new placer called *Sankeerna* was developed which is described in Section 4. The new placer *Sankeerna* is illustrated with an example in Section 5. The experimental setup to evaluate *Sankeerna* is described in Section 6. Results are tabulated and improvements obtained are discussed in Section 7. Conclusions of research work carried and future scope are given in Section 8.

II. RELATED WORK

Classical approaches to placement are reviewed in [10, 11, 3, 12, 13] and recent methods in [14, 15, 16]. The placement methods are classified based on the way the placement is constructed. Placements methods are either Constructive or Iterative [13]. In constructive method, once the components are placed, they will never be modified thereafter. An iterative method repeatedly modifies a feasible placement by changing the positions of one or more core cells and evaluates the result. Because of the complexity, the circuits are partitioned before placement. The constructive methods are (a) Partitioning-based which divide the circuit into two or more sub circuits [17] (b) Quadratic assignment which formulates the placement problem as a quadratic assignment problem [18, 19, 20] and (c) Cluster growth which places cells sequentially one by one in a partially completed layout using a criteria like number of cells connected to a already placed cell [21]. Main iterative methods are (a) Simulated annealing [22, 23,35], (b) Simulated evolution [15, 24] and (c) Force-directed [25, 20].

Another classification based on the placement technique used, was given in [20]. The placers were classified into three main categories namely (a) Stochastic placers which use simulated annealing which find global optimum with high CPU time, (b) Min-cut placers which recursively divide the netlist and chip area and (c). Analytical placers which define an analytical cost function and minimise it using numerical optimization methods. Some placers may use a combination of these techniques. These methods use only component cell dimensions and interconnection information and are not directly coupled to the synthesis.

The methods which use structural properties of the circuit are (a) Hierarchical placement [27] (b) Re-synthesis [9] and (c) Re-timing. There are algorithms which use signal flow and logic dependency during placement [28, 29]. In [28], critical paths are straightened after finding the zigzags. When placement is coupled with synthesis, this extra burden of finding criss-crosses is not required. In [30], using the structure of the interconnection graph, Placement is performed in a spiral topology around the centre of the cell array driven by a Depth First Search (DFS) on the interconnection graph. The algorithm has linear time complexity to the number of cells in the circuit.

To obtain delay optimized placements, timing driven placement methods are used [31, 32, 33, 34, 36, 37, 38, 39, 40]. The idea is to reduce the wire length on certain paths instead of total wire length. These methods are either path based or net based. The longest path delays are minimized in path based methods. Finding longest paths exponentially grows with the complexity of the design. Timing constraints are transformed into net-length constraints in the net based algorithms. Then a weighted wire length minimized placement is done iteratively until better timing is achieved. The drawbacks of

this method are (a) delay budgeting is done without physical placement feasibility and (b) it is iterative. At the end of the iteration, the solution produced is evaluated.

To control congestion and to achieve routability, white spaces are allocated at the time of placement [41, 26, 42, 43]. The problem with this approach is, it increases area which in turn will increase wire length and delay. In [44], it was shown that minimising wire length improves routability and layout quality. Allocating white space may not be the right approach to achieve routability. It is better to minimise the wire length instead of allocating white spaces. The white space allocated may not be the right place required for the router.

The studies in [45] have shown that existing placement algorithms produce significantly inferior results when compared with the estimated optimal solutions. The studies in [12] show that the results of leading placement tools from both industry and academia may be up to 50% to 150% away from optimal in total wire length.

The design flow convergence is another main requirement of Synergistic approach towards DSM design [4]. Placement plays a major role in this. As mentioned in [4], there are three types of design flows for Deep Sub Micron (DSM) namely (a) Logic synthesis drives DSM design (b) Physical design drives DSM design (c) Synergistic approach towards DSM design. In the last method (c), it is required to create iteration loops which tightly couple various levels of design flow. The unpredictability of area, delay, and routability of the circuits from synthesis to layout, is another major problem. The study in [46, 47] indicated that the non-convergence of design process is due to non-coupling of synthesis [48, 49, 50, 51] to placement process. We need a faster way of estimating area and delay from pre-place to post-route. If we fail to achieve this, we may not have a clue to converge.

From the above analysis of the state of the art placement algorithms, we feel that there is still scope for improvement and the need for better placement algorithms meeting the requirements as mentioned in section 3. In the next Section, we describe the basis for our new algorithms, which try to solve some of the problems mentioned above.

III. BASIS FOR THE NEW ALGORITHMS

The new placement algorithm should have the following features (a) linear or polynomial time complexity with respect to number of cells in the circuit, (b) awareness of synthesis and routing assumptions and expectations, that is, tight coupling of synthesis and routing as mentioned in [4, 46, 47], (c) achieving minimum area and delay, (d) produce routable layouts without Design Rule Check (DRC) violations, by proper wire planning during placement, (e) delay of final layout should be predictable with trial routes and (f) should smoothly interface with synthesis and routing tools.

In this section, we explain the basis for the *Sankeerna* algorithms. Since the circuits are to be placed to achieve minimum area and delay, we first try to find out what is the minimum area and delay which are achievable for a given circuit and technology library. The minimum area achievable is the sum of widths of all cells multiplied by the height of the standard cells. For a given standard cell technology library, the height is same for all cells.

To find out the minimum delay achievable, we use Figure 1(a) to explain the delay calculation process.

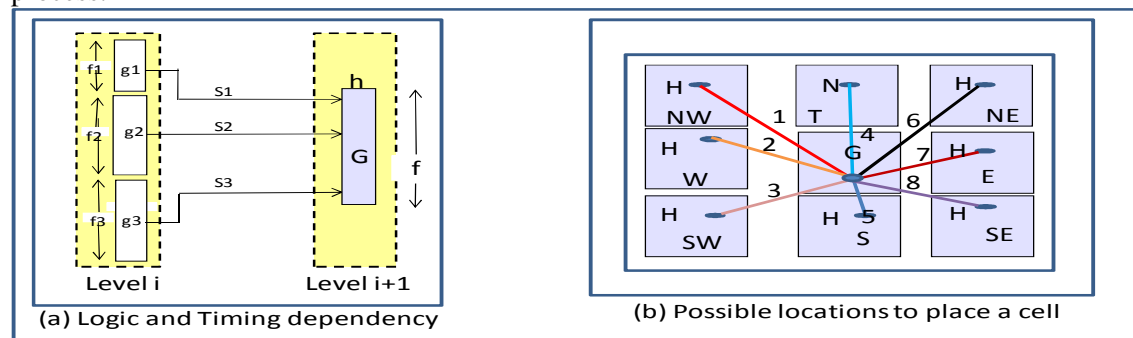


Figure 1 (a) Logic & timing dependency (b) possible locations to place a cell

There are 4 gates marked as g_1 , g_2 , g_3 and G . The gates g_1 , g_2 and g_3 are at level i and G is at level $i+1$ as per the logic of the circuit. Let d is delay of gate G , a is maximum of arrival times of (g_1 , g_2 , g_3) at the input of G , b is block delay of G , f is fan-out delay of G and w is wire delay of G . Then d is equal to sum of a , b , f and w . The d is dependent on the arrival times of inputs g_1 , g_2 and g_3 which are in transitive fan-in of G . The gates g_1 , g_2 and g_3 in turn will have inputs either from Primary Inputs (PIs) or from other gates. So delay of G is dependent on the transitive fan in of all gates connected to g_1 , g_2 and g_3 . To minimise d , the only thing that can be done at placement time is to minimize the wire delay w . Wire delay w depends on the wire lengths s_1 , s_2 and s_3 , and several other factors. We will consider only wire length here. To minimize wire lengths s_1 , s_2 and s_3 which are outputs of gates g_1 , g_2 and g_3 , they are to be placed at physically closest location to G . The possible places for a cell H which are nearer to a cell G are shown in Figure 1(b). H can be placed in any of the eight positions indicated by NW, N, NE, W, E, SW, S and SE. The distance from H output pin to G input pin for all these 8 possible locations depends on the width and height of H , G , and pin locations on these two cells. The lines 1, 2, 3, 4, 5, 6, 7 and 8 in Figure 1 show the Euclidean distance for all 8 positions. The same procedure can be adopted to calculate the Manhattan distance. The Physically Shortest Place (PSP) is the location which has minimum Manhattan distance. In real technology libraries, the cell positions are specified by a rectangle or a set of rectangles, not just as a point. So, we can choose to connect anywhere on the rectangle based on the closeness to the destination cell. Out of available locations, the one with the minimum Manhattan distance is on the Physically Shortest Path (PSP).

Let r be the required time of a gate, and a be the arrival time, then the slack s at gate G is r minus a . Out of all slacks of inputs to a gate G , the Worst Negative Slack (WNS) indicates that the cell is on the critical path. The inputs g_1 , g_2 and g_3 which are more critical are to be placed closer to gate G when compared to others. This argument has to be recursively applied to all gates which are in transitive fan in of g_1 , g_2 and g_3 . That is, placing g_1 nearer to G means, placing all the cells which are in transitive fan in of g_1 nearer to G . All gates which are in transitive fan in of g_1 are to be placed on PSP giving priority to cells which have higher WNS. Let WNS of g_1 , g_2 and g_3 be -2, -1 and -3 respectively. Placement priority is g_3 first, then g_1 and last g_2 . The minimum delay is achieved when g_1 , g_2 and g_3 are placed in PSP from Primary Outputs (POs) to Primary Inputs (PIs).

Sankeerna uses constructive method of placement. Starting from the Primary Output (PO), cells are placed on PSP as explained above. The height to width ratio of the smallest drive capability inverter is 4 for the standard cell library we have used for conducting experiments in this paper. So the row of a standard cell becomes the Physically Shortest Path (PSP). WNS at each node input, decides Delay-wise Shortest Path (DSP). *Sankeerna* combines PSP and DSP concepts explained above to produce routable placements minimizing area, delay and wire length in linear time. These concepts are further illustrated with a fully worked out example in Section 5. The next Section explains the algorithms used in *Sankeerna*.

IV. ALGORITHMS USED IN SANKEERNA

We have modified the algorithms of ANUPLACE [46] to produce delay optimized placements using a constructive method. *Sankeerna* reads the benchmark circuit which is in the form of a netlist, taken from "SIS" synthesizer [52], builds trees with Primary Outputs (POs) as roots. In SIS, we specify zero as the required time at the POs and print the worst slack at each node. This slack information is read along with the netlist into *Sankeerna*. The inputs are sorted based on this slack, with descending order of time criticality at each node. Starting from the root to the leaf node, nodes are placed on the layout after finding the closest free location. At each node, most time critical node is placed first using a modified Depth First Search (DFS) method. Priority is given to time when compared to depth of the tree. It was proved that placement of trees can be done in polynomial time [7]. A Depth First Search (DFS) algorithm was used in [30] which has linear time complexity to the number of connections. Tree based placement algorithms reported in literature have either linear time or $O(n \log n)$ time complexity [7, 53, 54, 55].

We have used benchmark circuits from SIS [52] synthesizer in “BLIF” format which are then converted into Bookshelf [56] format using converters provided in [59, 60, 41, 61]. The normal placement benchmark circuits [57, 45] are not useful because they give only cell dimensions and interconnect information. Timing, cell mapping, logic dependency and other circuit information from synthesizer are not available in these placement benchmarks. These converters do not use technology library information for cell dimensions and pin locations. *Sankeerna* reads the technology information consisting (a) cell names, (b) cell dimensions height and width, (c) pin locations on the cells, (d) timing information, and (e) input load from a file. Using this technology information, *Sankeerna* generates the benchmark circuit in bookshelf format with actual cell dimensions and pin locations. Once the trees are created the slack information is read into the tree data structure. *Sankeerna* algorithm is shown in Figure 2.



Figure 2 *Sankeerna* algorithms

As shown in the Figure 2, the “place_ckt” function places the trees pointed by Primary Outputs (POs) one after the other using “place_cell” function starting at row 1 at x=0 and y=0. The “place_cell” function works as follows.

- Place the cell pointed by root using “place_one_cell”.
- For each input, if it is a primary input, place it using “place_one_cell”, if not; call “place_cell” with this input recursively.

The “place_one_cell” function finds the best place to place the cell using “find_best_place” function and places the cell at this location. The “find_best_place” function works as follows. As the placement progresses, the “used space” and “available space” are marked, C is the current cell and the next cell will be placed closer to this cell wherever space is available. The current cell placement is depicted in Figure 3.

The “find_best_place” function checks the availability of space using “check_availability_on_row” function on the same row of C and on the rows above and below the current row of C. The possible places for a cell H which are nearer to a cell G are shown in Figure 3. Out of available locations, the one with the minimum distance from the parent cell is chosen. The cell is placed at this location. The “check_availability_on_row” function keeps two pointers x1 and x2 for each row. Initially x1 and x2 are initialised to zero before the start of placement. When this function is invoked, it gets these two pointers x1 and x2 from the table corresponding to this row. It then calculates whether in the available space, this cell of “width” can be placed such that the row width will be less than or equal to the layout width. If space is available, it returns x2 as the available location on this row.

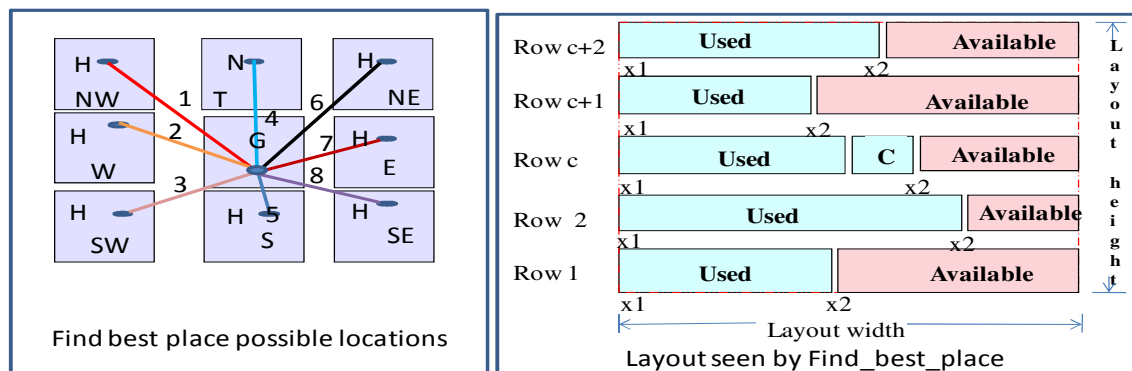


Figure 3 Find_best_place possible locations and the layout seen by Find_best_place

Complexity of the algorithm

If n is the number of the cells in the circuit, “place_one_cell” function is invoked n times. “Place_one_cell” calls “find_best_Place” which uses “check_availability_on_row”. So “find_best_Place” is executed n times. Each time, it calculates wire lengths for possible locations to choose the best one. “check_availability_on_row” performs one comparison operation. So number of operations is linearly proportional to n , after construction of the tree. So the complexity of algorithm is of the order of n .

V. ILLUSTRATION OF SANKEERNA WITH AN EXAMPLE

The algorithms used in *Sankeerna* are illustrated with an example whose logic equation is given below, taken from [49].

$$Y = ab\bar{e}g + abfg + ab\bar{e}g + ac\bar{e}g + acfg + ac\bar{e}g \\ + a\bar{e}g + dfg + d\bar{e}g + bh + bi + ch + ci$$

The logic diagram with technology mapped cells and the tree built by *Sankeerna* for the above logic diagram with the slacks are shown in Figure 4 and 5 along with the sequence of placement based on the time criticality.



Y Primary Output

1 [UY] 0.347891

2 [Ua356] 0.347891

22 [Ua358] 0.330634

19 [Ua349] 0.160878

3 [Ua347] 0.347891

23 [Ua434] 0.330634

25 [Ua444] 0.154191

21 [Ua442] 0.160878

20 [Ua432] 0.160878

17 [Ua446] 0.167536

18 [Ua445] 0.167536

4 [Ua360] 0.347891

24 [Ua433] 0.330634

26 [Ua443] 0.154191

5 [Ua431] 0.229631

6 [Ua351] 0.347891

15 [Ua430] 0.229631

12 [Ua364] 0.339816

7 [Ua362] 0.347891

13 [Ua439] 0.339816

14 [Ua438] 0.323096

8 [Ua438] 0.347891

9 [Ua441] 0.339262

11 [Ua435] 0.323096

10 [Ua437] 0.347891

10 [Ua440] 0.339262

Figure 5 Tree built by *Sankeerna* for the example



Figure 6 *Sankeerna* and Public Domain Placer (PDP) Placements of example

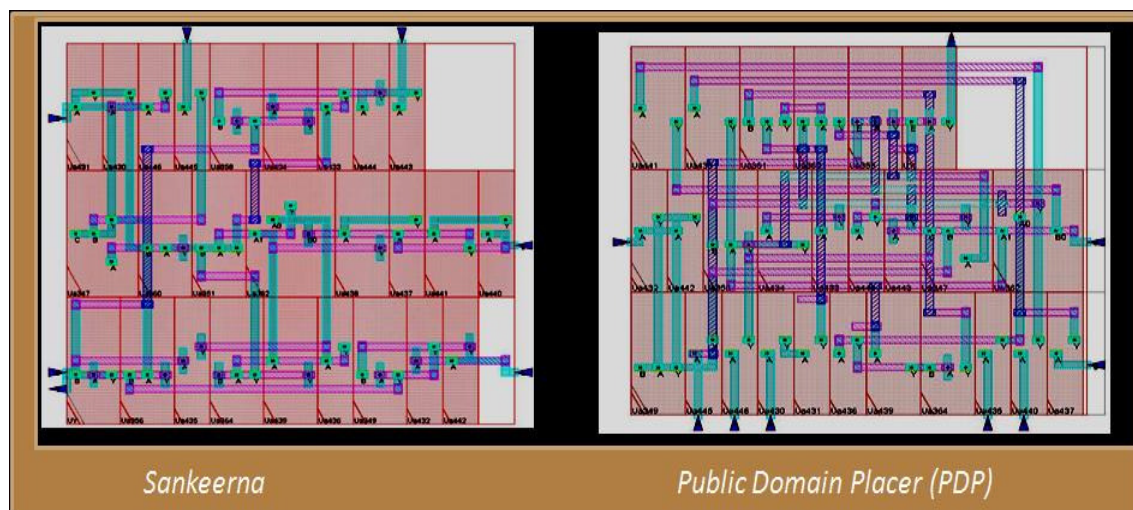


Figure 7 *Sankeerna* & PDP layout of example

The experimental set up to evaluate *Sankeerna* using benchmark circuits is explained in the next section.

VI. TEST SETUP

In this section, we describe the test set up used to evaluate *Sankeerna*. The test set up is shown in Figure 8.

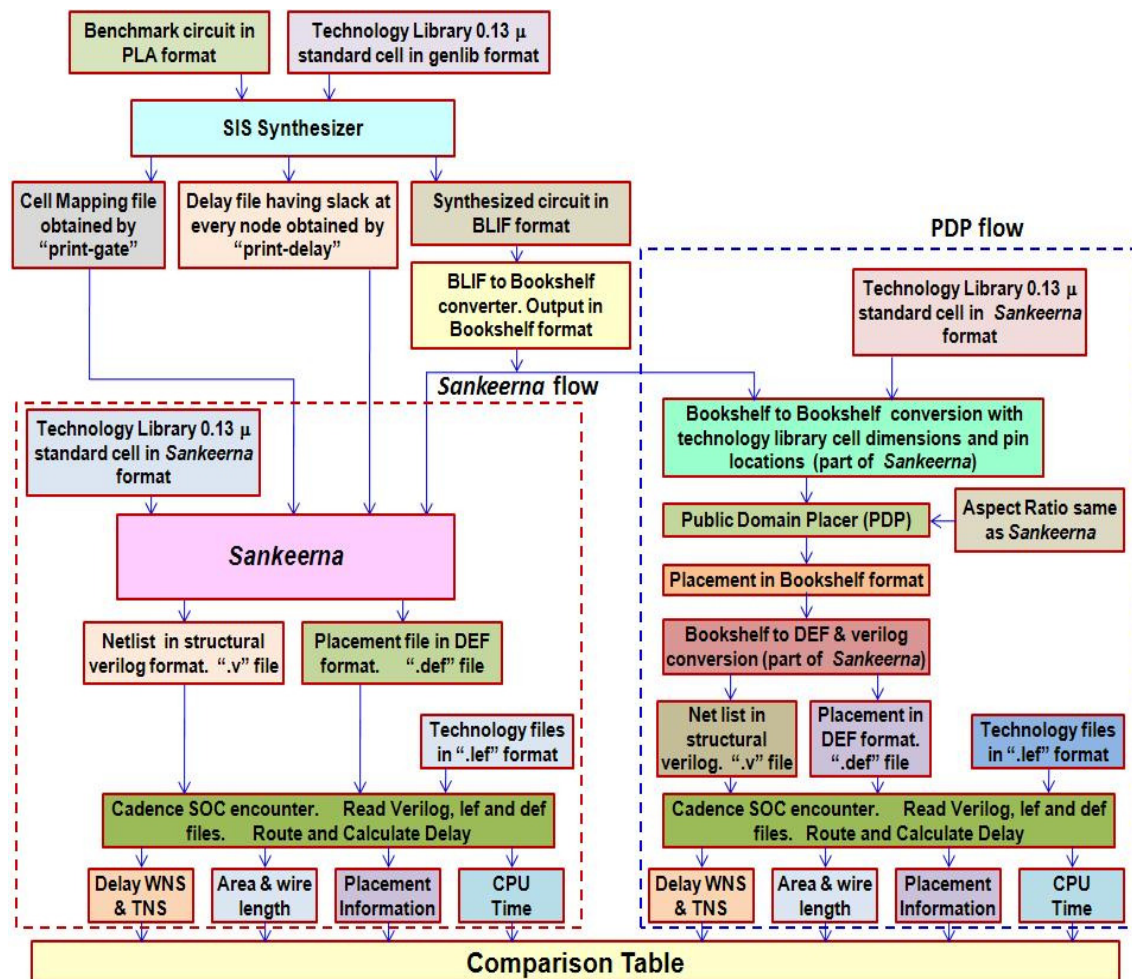


Figure 8 Test setup showing *Sankeerna* VLSI Design Flow

The benchmark circuits are taken in the form of a PLA. Some of them are from MCNC benchmarks. We also added few other circuits like multiplexers, adders and multipliers. We have used 0.13 micron, 8 metal layers, standard cell technology library for these experiments. Public domain SIS synthesizer [52] is used for synthesizing the benchmark circuits.

We created the library file required for SIS in genlib format using the information from the data sheets and “.lef” files of this particular technology. Three files namely (a) delay file (b) Cell mapping file (c) BLIF file are generated from SIS. The delay file consists of node name and slack at each node. Only block delays and fan out delays are considered. No wire delay or wire load models are used. This delay file is created using the information generated by “print_delay” command of SIS. The cell mapping file consists of node name and mapped library cell name. This file is created using the information generated by “print_gate” command of SIS. The BLIF file is created using “write_blif” command of SIS. The BLIF output is then converted into Bookshelf format using the public domain tools available at the web site [59, 60, 41, 61] using the utility “blif2book-Linux.exe filename.blif filename”. Using 0.13 micron standard cell technology files, *Sankeerna* generates a file in bookshelf format using the cell dimensions and pin locations of the given technology library. This file is used for placement by *Sankeerna* and also by Public Domain Placer (PDP) [59].

Bench marks are placed in case of PDP flow using “time LayoutGen-Lnx32.exe -f filename.aux -AR 1.5 -saveas filename” [59]. The width to height ratio is 3:2 which is same for *Sankeerna*. *Sankeerna* gives the placement output in “.def” format [62] (“.def file”). The mapped netlist is given out in the form of structural verilog file (“.v file”). Cadence® SOC Encounter® v06.10-p005_1 [58] is used for routing and calculating the delay of the placement produced. The verilog (.v file) and placement (.def file) are read into SOC Encounter®. The 0.13 micron standard cell technology files, consisting of

“.lef” files and timing library files “.tlf” are read into SOC encounter. We did a trial route and then detailed routing using Cadence NanoRoute® v06.10-p006 [40]. The delays Worst Negative Slack (WNS) and Total Negative Slack (TNS) “from all inputs to all outputs” are noted. The CPU time for detailed routing is also noted. Various other characteristics of placed circuits namely standard cell area, core area, wire length and Worst Negative Slack (WNS) are noted. Results for various benchmarks are shown in Table 1. We compared the results produced by *Sankeerna* with a Public Domain Placer (PDP). The PDP flow is also shown in Figure 8. The PDP uses the benchmark circuits in bookshelf format. The BLIF file generated by SIS is converted into bookshelf format with cell dimensions and pin locations as per the 0.13 micron standard cell library. We have used same aspect ratio for *Sankeerna* and PDP. The aspect ratio for these experiments was 0.4 for height and 0.6 for width. To have a fair comparison, the Primary Inputs and Primary Outputs are placed at the boundary of the core for both placements produced by *Sankeerna* and PDP. The output from PDP is in bookshelf format (.pl, .scl files). This is converted into “.def” format. The netlist is generated in the form of structural verilog file. All the utilities used to convert the above file formats are developed as software package of *Sankeerna*. The verilog, “.def”, the technology file (.lef file) and timing files (.tlf files) are read into Cadence® SOC Encounter® [58]. The detailed routing and delay calculations are carried out. A Linux machine with dual Intel® Pentium® 4 CPU @3.00GHz and 2 GB memory was used for running *Sankeerna* and PDP. For running SOC Encounter® [58], Sun Microsystems SPARC Enterprise® M8000 Server with 960 MHz CPU and 49 GB memory was used. The results are tabulated in Table 1.

VII. RESULTS AND DISCUSSION

The Table 1 shows the results of the placed circuits using existing Public Domain Placer (PDP) [59] and *Sankeerna*.

There is an average improvement of 24% in delay (Worst Negative Slack WNS), 8% in area and 75% in wire length after detailed routing with Nano-route of Cadence [58] when compared to Public Domain Placer (PDP). *Sankeerna* uses only 3% of area extra over the standard cell area as shown under “S” in Table 1 where as PDP uses 12%. In case of bigger bench marks, namely, alu4, e64, mul5510 and add889, there are thousands of DRC violations in case of PDP. This is due to increased usage of wire length. So those placements are not useful.

To compare *Sankeerna* with a commercial tool, we conducted the following experiment. We placed the benchmark circuit “alu4” using *Sankeerna* and did the detailed routing and delay calculation as mentioned earlier. Then the results given by *Sankeerna* are noted. We specified the same dimensions of width and height in the Commercial Tool Placer (CTP) after reading the verilog file into the tool. We then ran timing driven placement of commercial tool. We then carried out detailed routing. CPU time used for *Sankeerna* and Commercial Tool Placer (CTP) are given in Table 2. SOC Encounter® took 2:27 CPU time for detailed routing using Nanoroute® [58] for *Sankeerna* placement without any DRC violations. SOC Encounter® took 6:45 CPU time for detailed routing using Nanoroute for CTP’s timing driven placement with 144 DRC violations in Metal layer 1. The layouts produced for *Sankeerna* placement and CTP’s placement are shown in Figure 9. The black marks in CTP layouts in the Figure 9 are the DRC violations. These are shown separately in the same Figure (extreme right block). Since there are so many DRC violations, the layout is not useful.

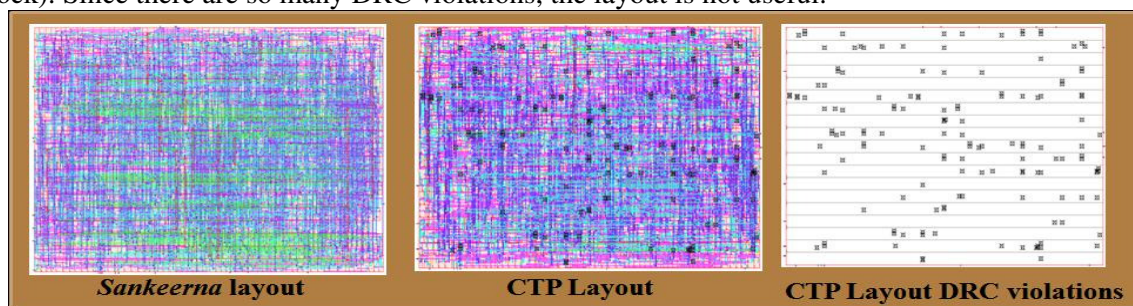


Figure 9 Layouts of ALU4 with *Sankeerna* & CTP

Table 1 Test Results

SI No	Name	Std Cell Area	Pre-place WNS	Sankeerna				PDP				% of area increase over Std Cell		% improvements over PDP		
				Area	WNS	Wire Length	DRC violations	Area	WNS	Wire Length	DRC violations					
												S	PDP	Ar ea	Wire Leng th	WN S
1	5xp1	912	-1.06	927	-1.274	2102	0	1013	-1.752	4037	0	2	11	8	92	38
2	9sym	1482	-1.19	1512	-1.568	5256	0	1647	-1.858	8125	0	2	11	8	55	18
3	alu4	6723	-2.41	6773	-3.237	33296	0	7470	-5.528	89520	13434	1	11	9	169	71
4	b12	536	-0.71	550	-0.850	1195	0	596	-0.891	1476	0	3	11	8	24	5
5	clip	1268	-1.51	1290	-1.762	3371	0	1409	-2.099	6170	0	2	11	8	83	19
6	cm82a	156	-0.54	173	-0.603	204	0	173	-0.598	233	0	11	11	0	14	-1
7	comp	1190	-1.24	1222	-1.467	2943	0	1322	-1.904	5391	0	3	11	8	83	30
8	con1	139	-0.42	148	-0.489	198	0	155	-0.513	249	0	6	11	5	26	5
9	cordic	732	-0.98	754	-1.152	1406	0	813	-1.250	2645	0	3	11	7	88	9
10	count	1275	-0.97	1290	-1.238	3535	0	1416	-1.466	5863	0	1	11	9	66	18
11	e64	7002	-0.96	7031	-1.657	44024	0	7780	-3.316	94434	17461	0	11	10	115	100
12	ex5	3057	-0.97	3133	-1.334	13181	0	3397	-1.899	23916	0	2	11	8	81	42
13	misex1	497	-0.76	509	-0.902	999	0	553	-0.967	1536	0	2	11	8	54	7
14	misex2	884	-0.64	927	-0.765	2284	0	983	-0.941	3475	0	5	11	6	52	23
15	mux8-1	175	-0.59	188	-0.655	224	0	194	-0.716	310	0	8	11	3	38	9
16	o64	1514	-0.68	1558	-0.816	2603	0	1682	-1.026	7562	0	3	11	7	191	26
17	rd53	341	-0.71	356	-0.827	640	0	379	-0.854	739	0	4	11	6	15	3
18	rd73	1044	-1.22	1059	-1.441	2715	0	1160	-1.834	4622	0	1	11	9	70	27
19	rd84	949	-1.06	974	-1.347	2513	0	1054	-1.702	4351	0	3	11	8	73	26
20	sao2	1008	-1.26	1059	-1.390	2564	0	1120	-1.792	4359	0	5	11	5	70	29
21	squar5	475	-0.85	492	-0.962	943	0	528	-0.995	1314	0	4	11	7	39	3
22	t481	217	-0.65	231	-0.699	228	0	241	-0.762	285	0	6	11	4	25	9
23	Z9sym	745	-1.23	784	-1.495	1976	0	828	-1.836	2952	0	5	11	5	49	23
24	example	115	-0.60	127	-0.651	137	0	168	-0.605	241	0	10	45	24	76	-7
25	add556	1516	-1.23	1558	-1.560	4737	0	1684	-1.984	7529	0	3	11	7	59	27
26	mul5510	8553	-3.02	8626	-4.765	44913	0	9504	-7.325	114480	17586	1	11	9	155	54
27	mul448	2161	-1.89	2203	-2.520	7008	0	2401	-3.039	14457	0	2	11	8	106	21
28	add667	2348	-1.55	2390	-2.140	8079	0	2608	-2.741	14472	0	2	11	8	79	28
29	add889	5708	-1.87	5742	-3.068	26907	0	6342	-4.172	62149	2398	1	11	9	131	36
30	add778	3422	-1.87	3486	-2.786	12462	0	3794	-3.294	22303	0	2	11	8	79	18
Average												3	12	8	75	24

Table 2 shows results of other benchmarks comparing with Commercial Tool timing driven Placer (CTP).

Table 2 Comparing Sankeerna with Commercial Tool Placer (CTP)

SI No	Name	Std Cell Area	Core Area	Floor Plan Dimension		Sankeerna				CTP			
						CPU Time	WNS	TNS	DRC violations	CPU Time	WNS	TNS	DRC violations
				Width	Height								
1	alu4	6723	6773	96.60	70.11	2:27	-3.24	-23.22	0	6:45	-3.35	-21.58	144
2	e64	7002	7031	100.28	70.11	8:21	-1.66	-93.93	0	12:57	-1.33	-73.01	127
3	mul5510	8553	8626	111.32	77.49	1:17	-4.77	-39.70	0	10:04	-4.47	-35.11	283
4	add889	5708	5742	91.54	62.73	2:42	-3.07	-21.97	0	10:08	-2.67	-19.29	103

For the same floor plan dimensions, CTP took more CPU time and could not produce final usable placement due to several DRC violations as shown in the last column of Table 2. *Sankeerna* placements are always routed without DRC violations, because wire planning is already done using shortest wires. This avoids iterations between placement and detailed routing steps. Even though, WNS and TNS values shown in the Table are comparable, because of DRC violations, the layout

produced by CTP is not useful; hence delay values have no meaning in case of CTP. So our method took much less time and produced better results when compared to Commercial Tool Placer (CTP). Pre-place delay is computed without taking wire delays into considerations. We have not used wire load models in SIS, because wire load models are not good estimates of final delay [2]. The post route delay includes delay due to wires and is the final delay of the circuit. We have compared pre-place delay with post route delay in Table 3.

Table 3 Pre-place delay versus post routed delay comparison

SI No	Name	Preplace	Sankeerna	PDP	Sankeerna	PDP
		WNS	WNS	WNS	WNS Diff %	WNS Diff %
1	alu4	-2.41	-3.36	-5.50	39.39	128.50
2	e64	-0.96	-1.72	-3.30	78.61	242.99
3	ex5	-0.97	-1.32	-1.85	36.37	91.19
4	mul5510	-3.02	-4.77	-6.84	57.57	126.12
5	mul448	-1.89	-2.52	-3.06	33.69	62.33
6	add889	-1.87	-3.07	-4.13	64.24	120.82
7	add778	-1.87	-2.79	-3.41	49.22	82.65
					51.30	122.09

The percentage difference of Worst Negative Slack of *Sankeerna* placements is much less (51.30% versus 122.09%) when compared to Public Domain Placer's (PDP's) values. This is due to the fact that PDP uses more wire when compared to *Sankeerna*.

Cadence® SOC Encounter® [58] has the facility called "Trial route", which performs quick global and detailed routing creating actual wires, estimating routing related congestion and capacitance values. We did trial route of the benchmarks and noted wire lengths, WNS and TNS values for *Sankeerna* and PDP placements. The results are compared in Table 4.

Table 4 Trial route Versus Detail (Nano) route, wire length and delay comparison

SI No	Name	Trial Route				Nano Route				Sankeerna Trial Vs Nano		PDP Trial vs Nano	
		Sankeerna		PDP		Sankeerna		PDP		Wire-length Diff %	WNS Diff %	Wire Length Diff %	WNS Diff %
		Wire-length	WNS	Wire length	WNS	Wire length	WNS	Wire length	WNS				
1	alu4	33418	-3.26	73683	-4.90	33296	-3.24	89520	-5.53	-0.36	-0.61	21.49	12.82
2	e64	43530	-1.69	79155	-2.80	44024	-1.66	94434	-3.32	1.14	-2.18	19.30	18.43
3	ex5	13486	-1.30	23531	-1.90	13181	-1.33	23916	-1.90	-2.27	2.38	1.64	-0.05
4	mul5510	44440	-4.63	97929	-6.30	44913	-4.77	114480	-7.33	1.06	3.03	16.90	16.27
5	mul448	7147	-2.44	14442	-3.03	7008	-2.52	14457	-3.04	-1.94	3.11	0.10	0.20
6	add889	27108	-3.01	54723	-4.03	26907	-3.07	62149	-4.17	-0.74	1.83	13.57	3.65
7	add778	12798	-2.77	22371	-3.26	12462	-2.79	22303	-3.3	-2.62	0.76	-0.30	1.10
Average %										-0.82	1.19	10.39	7.49

The percentage differences from trial route to detailed (Nano) route are shown in the last 4 columns of Table 4 for both *Sankeerna* and PDP. There is decrease in wire length by 0.82%, 1.19% increase in WNS in case of *Sankeerna*. The PDP placements took 10.39% more wire, WNS increased by 7.49%. So the Trial Route produced delays are good estimates in case of *Sankeerna* when compared to PDP placements. So we can get quick estimate of delay for *Sankeerna* produced placements. All these advantages add towards tighter coupling of VLSI design flow as envisaged in [4] to evolve Synergistic Design Flow. *Sankeerna* scales well as the complexity increases, because we are using constructive placement. The placement was found to converge after routing in all the test cases we have tried, that is, layout generated was without DRC violations after routing. To prove this point, we have taken bigger benchmarks out of Table 1 and improvements obtained are shown in Table 5.

As can be seen from the Table 5, *Sankeerna* took only 1.3% extra area over the standard cell area where as PDP took 11% extra area. The 1.3% extra area is left at the right edges of all rows. This is the bare minimum area required to make the placement possible. The area improvement of *Sankeerna* over PDP is 8.8%. The most interesting fact is that wire length is 114.4% more for PDP when compared to *Sankeerna*. WNS improved by 46.2%. There are several DRC violations after detailed routing in case of PDP. Hence those layouts are not useful. *Sankeerna* used bare minimum area and produced better timings which were routable.

Table 5 As complexity increases, *Sankeerna* performs better over PDP.

Sl No	Name	Std Cell Area	Pre-place WNS	<i>Sankeerna</i>				PDP				% Area Increase over Std Cell area		% Improvements		
				Area	WNS	Wire Length	DRC violations	Area	WNS	Wire Length	DRC violations	S	PDP	Area	Wire Length	WNS
1	alu4	6723	-2.41	6773	-3.237	33296	0	7470	-5.528	89520	13434	0.7	11.1	9.3	168.9	70.8
2	e64	7002	-0.96	7031	-1.657	44024	0	7780	-3.316	94434	17461	0.0	11.0	9.6	114.5	100.1
3	ex5	3057	-0.97	3133	-1.334	13181	0	3397	-1.899	23916	0	2.0	11.0	7.8	81.4	42.4
4	mul551	8553	-3.02	8626	-4.765	44913	0	9504	-7.325	114480	17586	1.0	11.0	9.2	154.9	53.7
5	mul448	2161	-1.89	2203	-2.520	7008	0	2401	-3.039	14457	0	2.0	11.0	8.2	106.3	20.6
6	add667	2348	-1.55	2390	-2.140	8079	0	2608	-2.741	14472	0	2.0	11.0	8.4	79.1	28.1
7	add889	5708	-1.87	5742	-3.068	26907	0	6342	-4.172	62149	2398	1.0	11.0	9.5	131.0	36.0
8	add778	3422	-1.87	3486	-2.786	12462	0	3794	-3.294	22303	0	2.0	11.0	8.1	79.0	18.2
Average												1.3	11.0	8.8	114.4	46.2

Area, wire length and delay are interrelated. If we use minimum area, as is the case with *Sankeerna*, we require less length wire. This in turn leads to minimum delay. So using *Sankeerna* flow avoids design iterations, because the flow is tightly coupled from synthesis to layout. Thus *Sankeerna* produced compact layouts which are better in delay when compared to Public Domain Placer (PDP) and Commercial Tool timing driven Placer (CTP).

VIII. CONCLUSIONS AND FUTURE SCOPE

We now summarise the features of *Sankeerna* for fitting it into a Synergistic Design Flow (SDF) as mentioned in Section 3.

The first requirement was (a) linear time complexity with respect to number of cells in the circuit. The *Sankeerna* placement algorithms have linear time complexity after construction of the tree. In a tightly coupled Synergistic Design Flow (SDF), trees are already built by synthesizer which can be directly used by the placer. So there is no need to construct them separately. Due to linear time complexity, *Sankeerna* scales well as the circuits become bigger.

The second requirement was (b) awareness of synthesis and routing assumptions and expectations, that is, tight coupling of synthesis and routing as mentioned in [4]. We have used logic dependency and timing information from synthesizer. Using this information, it properly guides the placement as mentioned in [46, 28, 29, 30, 47]. As shown in Table 3, Pre-place to post routed delay variation for *Sankeerna* was 51.30% when compared to 122.09% for Public Domain Placer (PDP). The values vary from 39.39% to 78.6% for *Sankeerna*. Where as the variation for PDP was, from 62.33% to 242.99% based on the circuit. So *Sankeerna* is more coupled to synthesizer's estimates when compared to PDP. *Sankeerna* placements were always routed without DRC violations as shown in Table 1 and 2. Where as PDP has thousands of violations for bigger circuits even after using 11% extra space when compared to *Sankeerna* which used only 0% to 1% extra over the area calculated by the synthesizer, which is bare minimum over the standard cell area. For the same floor plan dimensions, Commercial Tool's timing driven Placer (CTP) produced hundreds of DRC violations as shown in Table 2 when compared to zero DRC violations for *Sankeerna*. In *Sankeerna* routability is achieved without white space allocation, because placements produced by *Sankeerna* use minimum length wires. As mentioned in Section (2) white space allocation increases area which in turn increase wire length. In conclusion, the placements produced by *Sankeerna* were always routable because it uses minimum wire when compared to PDP and CTP.

The third requirement was (c) achieving minimum area and delay. Area increase was only 1.3% over the standard cell area which is calculated by the synthesizer. This value for PDP was 11%. As shown in Table 5, *Sankeerna* performed better as the complexity increased when compared to Public Domain Placer (PDP). Wire length improved by 114.4% and delay by 46.2% when compared to PDP. The fourth requirement was (d) placement produced should be routable without Design Rule Check (DRC) violations, that is, wire planning has to be done during placement. As shown in Table 5, PDP could not produce usable placements due congestion and resulted in thousands of DRC violations in four cases out of 8 test cases. So design flow did not converge in case of PDP. It is unpredictable, because, in some cases it converged. In case of *Sankeerna*, it always converged and convergence is

predictable due to minimum wire length and proper wire planning done during placement. As shown in Table 2, same non-convergence and unpredictability was noticed in case of Commercial Tool's Placer (CTP).

The fifth requirement was (e) delay of final layout should be predictable with trial routes. As shown in Table 4, for *Sankeerna*, wire length decreased by 0.82% from trial route to detailed route, where as it increased by 10.39 % for PDP. The delay increase was 1.19% for *Sankeerna*, where as it is 7.49% for PDP. Thus, the wire planning done by *Sankeerna* was maintained after routing, where as it varied in the case of PDP. Trial route was good estimate for *Sankeerna*. So there is convergence and tight coupling between placements and routing in case of *Sankeerna*.

The sixth requirement was (f) placer should smoothly interface with synthesis and routing tools. As shown in Figure 8, the test set up showing *Sankeerna* design flow, *Sankeerna* smoothly interfaces with the existing synthesizer and router. We have not used any wire load models during synthesis as it was demonstrated that they were not useful [2]. The slack at each node and cell mapping information were already available with the synthesizer. The router was interfaced through verilog netlist and "def" [62] file for placement information. Pre-place and trial route delay calculations were done by the commercial tool [58], which were good estimators in case of *Sankeerna* for a real standard cell technology library. As can be seen from the experiments, there were no design iterations among synthesis, placement and routing in case of *Sankeerna* to achieve the results shown.

Area, wire length and delay calculations of Pre-place, trial and post route were done by the Commercial tool. This validates that there is no error in measuring these values while conducting these experiments.

The features and effectiveness comparison of *Sankeerna* with other published placement techniques is elaborated here. In *Sankeerna* the cells which are logically dependent are placed closer to each other [29], whereas in other placement algorithms the cells are randomly scattered and create zigzags and criss-crosses that leads to increase in congestion, wire length and delays. The random scattering of the cells even leads to the unpredictability in the final layout that result into non-convergent iterations. This will result in non-convergent iterations. Because wires are shorter in our placement and wires are planned by selecting closest locations during placement, congestion is less and detailed routing always gets completed using minimum area for wires. This automatically leads to minimum delays. The most critical paths automatically get higher priority, without going in for path based placement, which grows exponentially with circuit complexity and is computationally expensive. As it can be seen from Figure 5, *Sankeerna* first establishes the most critical path and rest of the logic is placed around it based on the logic dependency. This is similar to the denser path placement of [30]. So the most critical path is placed, along physically and delay wise shortest path as mentioned in Section 3. Since our method is constructive, it scales well for bigger circuits. We are planning to test with bigger test cases in future. The Circuit is naturally partitioned when trees are built, rooted by Primary Outputs (POs) by *Sankeerna*. So there is no additional burden of extracting cones as in [27, 29] or partitioning the circuit as is the case in most of the placers. Global signal flow is kept in mind all through the placement by using Depth First Search (DFS) for all sub trees rooted at various levels of logic, unlike other placement methods, which randomly scatter the cells. Trial route can be used for quick estimate of delay which will be good estimate in case of *Sankeerna* as explained earlier. As mentioned in [2], using wire load models misleads whole design process resulting in non-convergence. So *Sankeerna* flow does not use wire load models. *Sankeerna* flow is always convergent and tightly coupled, which gives estimates of area, wire length and delay using existing layout tools like Trial route of Cadence's SOC Encounter® [58], which are not far away from the values obtained after detailed routing. Thus *Sankeerna* approach is useful towards evolving Synergistic Design Flow (SDF), which is to create iteration loops that are tightly coupled at the various levels of design flow as mentioned in [4].

ACKNOWLEDGEMENTS

We thank Dr. K.D. Nayak who permitted and guided this work to be carried out in ANURAG. We also thank members of ANURAG who reviewed the manuscript. Thanks are due to Mrs. D. Manikyamma and Mr. D. Madhusudhan Reddy for preparation of the manuscript.

REFERENCES

- [1] Randal E. Bryant, et al., (2001), "*Limitations and Challenges of Computer-Aided Design Technology for CMOS VLSI*", *Proceedings of the IEEE*, Vol. 89, No. 3, pp 341-65.
- [2] Gosti, W., et al., (2001), "*Addressing the Timing Closure Problem by Integrating Logic Optimization and Placement*", *ICCAD 2001 Proceedings of the 2001 IEEE/ACM International Conference on Computer-aided design*, San Jose, California , pp 224-231.
- [3] Shahookar K & Mazumder P, (1991), "*VLSI cell placement techniques*" *ACM Computing Surveys*, Vol. 23, No. 2.
- [4] Kurt Keutzer., et al., (1997), "*The future of logic synthesis and physical design in deep-submicron process geometries*", *ISPD '97 Proceedings of the international symposium on Physical design*, ACM New York, NY, USA, pp 218-224.
- [5] Coudert, O, (2002), "*Timing and design closure in physical design flows*", *Proceedings. International Symposium on Quality Electronic Design (ISQED '02)*, pp 511 – 516.
- [6] Wilsin Gosti , et al., (1998), "*Wireplanning in logic synthesis*", *Proceedings of the IEEE/ACM international conference on Computer-aided design*, San Jose, California, USA, pp 26-33.
- [7] Yifang Liu, et al., (2011), "*Simultaneous Technology Mapping and Placement for Delay Minimization*", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 30 No. 3, pp 416–426.
- [8] Pedram, M. & Bhat, N, (1991), "*Layout driven technology mapping*", *28th ACM/IEEE Design Automation Conference*, pp 99 – 105.
- [9] Salek, A.H., et al., (1999), "*An Integrated Logical and Physical Design Flow for Deep Submicron Circuits*", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 9, pp 1305–1315.
- [10] Naveed A. Sherwani, (1995), "*Algorithms for VLSI Physical Design Automation*", Kluwer Academic Publishers, Norwell, MA, USA.
- [11] Sarrafzadeh, M., & Wong, C.K., (1996), "*An introduction to VLSI Physical Design*", The McGraw-Hill Companies, New York.
- [12] Jason Cong, et al., (2005), "*Large scale Circuit Placement*", *ACM Transactions on Design Automation of Electronic Systems*, Vol. 10, No. 2, pp 389-430.
- [13] Yih-Chih Chou & Young-Long Lin, (2001), "*Performance-Driven Placement of Multi-Million-Gate Circuits*", *ASICON 2001 Proceedings of 4th International Conference on ASIC*, Shanghai, China, pp 1-11.
- [14] Yao-Wen Chang, Zhe-Wei Jiang and Tung-Chieh Chen, (2009), "*Essential Issues in Analytical Placement Algorithms*", *Information and Media Technologies*, Vol. 4, No. 4, pp.815-836
- [15] Bunglowala, A. & Singhi, B.M., (2008), "*Performance Evaluation and Comparison and Improvement of Standard Cell Placement Techniques in VLSI Design*", *First international conference on Emerging Trends in Engineering and Technology*, Nagpur, Maharashtra, 468 – 473.
- [16] B. Sekhara Babu, et al (2011), "*Comparison of Hierarchical Mixed-Size Placement Algorithms for VLSI Physical Synthesis*", *CSNT '11 Proceedings of the 2011 International Conference on Communication Systems and Network Technologies*, IEEE Computer Society Washington, DC, USA, pp 430-435.
- [17] Mehmet Can Yildiz & Patrick H. Madden, (2001), "*Global objectives for standard cell placement*", *GLSVLSI '01 Proceedings of the 11th Great Lakes symposium on VLSI*, ACM New York, NY, USA, pp 68 – 72.
- [18] C. J. Alpert, et al., (1997), "*Quadratic Placement Revisited*", *34th ACM/IEEE Design Automation Conference*, Anaheim, pp 752-757.
- [19] N. Viswanathan et al., (2007), "*FastPlace 3.0: A Fast Multilevel Quadratic Placement Algorithm with Placement Congestion Control*", *ASP-DAC '07 Proceedings of the 2007 Asia and South Pacific Design Automation Conference*, IEEE Computer Society Washington, DC, USA, pp 135-140.
- [20] Spindler, P, et al., (2008), "*Kraftwerk2—A Fast Force-Directed Quadratic Placement Approach Using an Accurate Net Model*", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 27 no. 8, 1398 – 1411.
- [21] Rexford D. Newbould & Jo Dale Carothers , (2003), "*Cluster growth revisited: fast, mixed-signal placement of blocks and gates*", *Southwest Symposium on Mixed Signal Design*, pp 243 – 248.
- [22] Carl Sechen & Alberto Sangiovanni-Vincentelli, (1985), "*The TimberWolf Placement and Routing Package*", *IEEE Journal of Solid-State Circuits*, vol. SC-20, No. 2, pp 510-522.

- [23] Khorgade, M. et al., (2009), "Optimization of Cost Function with Cell Library Placement of VLSI Circuits Using Simulated Annealing", 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET), Nagpur, India, pp 173 – 178.
- [24] Yoshikawa, M. & Terai, H., (2005), "A GA-based timing-driven placement technique", Sixth international conference on Computational Intelligence and Multimedia Applications, pp 74 – 79.
- [25] Andrew Kennings & Kristofer P. Vorwerk, (2006), "Force-Directed Methods for Generic Placement", IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems, Vol. 25, NO. 10, pp 2076-2087.
- [26] Chen Li et al., (2007), "Routability-Driven Placement and White Space Allocation", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Volume: 26 Issue:5, pp 858 – 871.
- [27] Yu-Wen Tsay, et al., (1993), "A Cell Placement Procedure that Utilizes Circuit Structural Properties", Proceedings of the European Conference on Design Automation, pp 189-193.
- [28] Chanseok Hwang & Massoud Pedram, (2006), "Timing-Driven Placement Based on Monotone Cell Ordering Constraints", Proceedings of the 2006 Conference on Asia South Pacific Design Automation: ASP-DAC 2006, Yokohama, Japan, pp 201-206.
- [29] Cong, J. & Xu, D, (1995), "Exploiting signal flow and logic dependency in standard cell placement", Proceedings of the Asian and South Pacific Design Automation Conference, pp 399 – 404.
- [30] Ioannis Fudos et al, (2008), "Placement and Routing in Computer Aided Design of Standard Cell Arrays by Exploiting the Structure of the Interconnection Graph", Computer-Aided Design & Applications, CAD Solutions, LLC, Canada, <http://www.cadanda.com/>, Vol. 5(1-4), pp 325-337.
- [31] Andrew B. Kahng & Qinke Wang, (2004), "An analytic placer for mixed-size placement and timing-driven placement", Proceedings of International Conference on Computer Aided Design, pp 565-572.
- [32] Jun Cheng Chi, et al., (2003), "A New Timing Driven Standard Cell Placement Algorithm", Proceedings of International Symposium on VLSI Technology, Systems and Applications, pp 184-187.
- [33] Swartz, W., & Sechen, C., (1995), "Timing Driven Placement for Large Standard Cell Circuits", Proc. ACM/IEEE Design Automation Conference, pp 211-215.
- [34] Tao Luo, et al., (2006), "A New LP Based Incremental Timing Driven Placement for High Performance Designs", DAC '06 Proceedings of the 43rd Annual Design Automation Conference, ACM New York, NY, USA, pp 1115-1120.
- [35] Wern-Jieh Sun & Carl Sechen, (1995), "Efficient and effective placement for very large circuits", IEEE Transactions on CAD of Integrated Circuits and Systems, Vol. 14 No. 3, pp 349-359.
- [36] Marek-Sadowska, M. & Lin, S.P. (1989), "Timing driven placement" IEEE international conference on Computer-Aided Design, Santa Clara, CA, USA, pp 94 - 97
- [37] Wilm E. Donath, et al., (1990), "Timing driven placement using complete path delays", Proceedings of the 27th ACM/IEEE Design Automation Conference, ACM New York, NY, USA, pp 84 – 89.
- [38] Bill Halpin, et al, (2001), "Timing driven placement using physical net constraints", Proceedings of the 38th annual Design Automation Conference, ACM New York, NY, USA, pp 780 – 783.
- [39] Xiaojian Yang, et al., (2002), "Timing-driven placement using design hierarchy guided constraint generation" Proceedings of the 2002 IEEE/ACM international conference on Computer-aided design, ACM New York, NY, USA, pp 177-180.
- [40] Riess, B.M. & Ettelt, G.G, (1995), "SPEED: fast and efficient timing driven placement", IEEE International Symposium on Circuits and Systems, Seattle, WA, USA, vol.1, pp 377 – 380.
- [41] Saurabh Adya and Igor Markov, (2003), "On Whitespace and Stability in Mixed-size Placement and Physical Synthesis", International Conference on Computer Aided Design (ICCAD), San Jose, pp 311-318.
- [42] Taraneh Taghavi, et al, (2006), "Dragon2006: Blockage-Aware Congestion-Controlling Mixed-Size Placer", ISPD '06 Proceedings of the 2006 international symposium on Physical design, ACM New York, NY, USA, pp 209 – 211.
- [43] Yi-Lin Chuang, et al., (2010), "Design-hierarchy aware mixed-size placement for routability optimization", IEEE/ACM International Conference on Computer-Aided Design (ICCAD), San Jose, CA, USA, pp 663 – 668.
- [44] Xiaojian Yang et al., (2002), "A standard-cell placement tool for designs with high row utilization", Proceedings of IEEE International Conference on Computer Design: VLSI in Computers and Processors, pp 45 – 47.
- [45] C. Chang, J. Cong, et al., (2004), "Optimality and Scalability Study of Existing Placement Algorithms", IEEE Transactions on Computer-Aided Design, Vol.23, No.4, pp.537 – 549.

- [46] Santeppa Kambham & Krishna Prasad K.S.R, (2011), "*ANUPLACE: A Synthesis Aware VLSI Placer to minimize timing closure*", *International Journal of Advances in Engineering & Technology*, Vol. 1, Issue 5, pp. 96-108.
- [47] Santeppa Kambham et al., (2008), "*New VLSI Placement algorithms to minimize timing closure problem*", *International conference on Emerging Microelectronics and Interconnection Technology (EMIT-08)*, IMAPS, Bangalore, India.
- [48] Brayton R K, et al., (1990), "*Multilevel Logic Synthesis*", *Proceedings of the IEEE*, Vol. 78, No. 2, pp-264-300.
- [49] Brayton R K, et al.,(1987), "*MIS: A Multiple-Level Logic Optimization System*", *IEEE Transactions on Computer Aided Design*, Vol.6, No.6, pp-1062-1081.
- [50] Rajeev Murgai, et al.,(1995), "*Decomposition of logic functions for minimum transition activity*", *EDTC '95 Proceedings of the European conference on Design and Test*, pp 404-410.
- [51] Fujita, M. & Murgai, R, (1997), "*Delay estimation and optimization of logic circuits: a survey*", *Proceedings of Asia and South Pacific Design Automation Conference*, Chiba, Japan, pp 25 – 30.
- [52] Sentovich, E.M., et al., (1992), "*SIS: A System for Sequential Circuit Synthesis*", Memorandum No. UCB/ERL M92/41, Electronics Research Laboratory, University of California, Berkeley, CA 94720.
- [53] M. Fischer & M. Paterson, (1980), "*Optimal tree layout (preliminary version,*", *STOC '80 Proceedings of the twelfth annual ACM symposium on Theory of computing*, ACM New York, NY, USA, pp. 177–189.
- [54] S. Chatterjee, et al., (2007) "*A linear time algorithm for optimum tree placement*", *Proceedings of International Workshop on Logic and Synthesis*, San Diego, California, USA
- [55] M. Yannakakis, "*A polynomial algorithm for the min-cut linear arrangement of trees*", *Journal of ACM*, vol. 32, no. 4, pp. 950–988, 1985.
- [56] Andrew Caldwell, et al., (1999), "*Generic Hypergraph Formats, rev. 1.1*", from <http://vlsicad.ucsd.edu/GSRC/bookshelf/Slots/Fundamental/HGraph/HGraph1.1.html>.
- [57] Jason Cong, et al, (2007), "*UCLA Optimality Study Project*", from <http://cadlab.cs.ucla.edu/~pubbench/>.
- [58] Cadence®, (2006), "*Encounter® Menu Reference*", Product Version 6.1, Cadence® Design Systems, Inc., San Jose, CA, USA. *Encounter® is the trademark of Cadence Design Systems, Inc., San Jose, CA, USA.
- [59] Saurabh Adya & Igor Markov, (2005), "*Executable Placement Utilities*" from <http://vlsicad.eecs.umich.edu/BK/PlaceUtils/bin>.
- [60] Saurabh N. Adya, et al., (2003), "*Benchmarking For Large-scale Placement and Beyond*", *International Symposium on Physical Design (ISPD)*, Monterey, CA, pp. 95-103.
- [61] Saurabh Adya and Igor Markov, (2002), "*Consistent Placement of Macro-Blocks using Floorplanning and Standard-Cell Placement*", *International Symposium of Physical Design (ISPD)*, San Diego, pp.12-17.
- [62] Cadence®, (2004), "*LEF/DEF Language Reference*", Product Version 5.6, Cadence® Design Systems, Inc., San Jose, CA, USA.

AUTHORS

Santeppa Kambham obtained B.Tech. in Electronics and Communication engineering from J N T U and M Sc (Engg) in Computer Science and Automation (CSA) from Indian Institute of Science, Bangalore. He worked in Vikram Sarabhai Space Centre, Trivandrum from 1982 to 1988 in the field of microprocessor based real-time computer design. From 1988 onwards, he has been working in the field of VLSI design at ANURAG, Hyderabad. He received DRDO Technology Award in 1996, National Science Day Award in 2001 and "Scientist of the Year Award" in 2002. He is a Fellow of IETE and a Member of IMAPS and ASI. A patent has been granted to him for the invention of a floating point processor device for high speed floating point arithmetic operations in April 2002.



Siva Rama Krishna Prasad Kolli received B.Sc. degree from Andhra University, DMIT in electronics from MIT, M.Tech. in Electronics and Instrumentation from Regional Engineering College, Warangal and PhD from Indian Institute of Technology, Bombay. He is currently working as Professor at Electronics and Communication Engineering Department, National Institute of Technology, Warangal. His research interests include analog and mixed signal IC design, biomedical signal processing and image processing.



A NEW VARIANT OF SUBSET-SUM CRYPTOSYSTEM OVER RSA

Sonal Sharma¹, Saroj Hiranwal², Prashant Sharma³

^{1&3}M Tech Student, Sri Balaji College of Engineering & Technology, Jaipur, Rajasthan

²Reader (CSE Dept), Sri Balaji College of Engineering & Technology, Jaipur, Rajasthan

ABSTRACT

RSA is an algorithm for public-key cryptography that is based on the presumed difficulty of factoring large integers, the factoring problem. RSA stands for Ron Rivest, Adi Shamir and Leonard, who first publicly described it in 1978. A user of RSA creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. The prime factors must be kept secret. In RSA if one can factor modulus into its prime numbers then the private key is also detected and hence the security of the cryptosystem is broken. The Subset-Sum cryptosystem (Knapsack Cryptosystem) is also an asymmetric cryptographic technique. The Merkle-Hellman system is based on the subset sum problem (a special case of the knapsack problem): An instance of the Subset Sum problem is a pair (S, t) , where $S = \{x_1, x_2, \dots, x_n\}$ is a set of positive integers and t (the target) is a positive integer. The decision problem asks for a subset of S whose sum is as large as possible, but not larger than t . This problem is NP-complete. However, if the set of numbers (called the knapsack) is super increasing, that is, each element of the set is greater than the sum of all the numbers before it; the problem is easy and solvable in polynomial time with a simple greedy algorithm. So in this paper we present a new algorithm (Modified Subset-Sum cryptosystem over RSA) which is secure against Mathematical attack, Brute-force attack, Factorization attack and Chosen-cipher-text attack on RSA as well as Shamir attacks. This paper also presents comparison between Modified Subset - Sum Cryptosystem and RSA cryptosystems in respect of security and performance.

KEYWORDS: Cryptography, Subset Sum, Public key, Private Key, RSA, Merkle-Hellman, Super increasing and Complexity

I. INTRODUCTION

To solve the problem of secure key management of Symmetric key cryptography, Diffie and Hellman introduced a new approach to cryptography and, in effect, challenged cryptologists to come up with a cryptographic algorithm that met the requirements for public-key systems. Public key cryptography uses a pair of related keys, one for encryption and other for decryption. One key, which is called the private key, is kept secret and other one known as public key is disclosed and this eliminate the need for the sender and the receiver to share secret key. The only requirement is that public keys are associated with the users in a trusted (authenticated) manner through a public key infrastructure (PKI). The public key cryptosystems are the most popular, due to both confidentiality and authentication facilities [1]. The message is encrypted with public key and can only be decrypted by using the private key. So, the encrypted message cannot be decrypted by anyone who knows only the public key and thus secure communication is possible.

In a public-key cryptosystem, the private key is always linked mathematically to the public key. Therefore, it is always possible to attack a public-key system by deriving the private key from the public key. The defense against this is to make the problem of deriving the private key from the public key as difficult as possible. Some public-key cryptosystems are designed such that deriving the

private key from the public key requires the attacker to factor a large number. The *Rivest-Shamir-Adleman (RSA)* and *Subset-Sum (Knapsack)* public key cryptosystems [2] are the best known examples of such a system. This paper presents a hybrid cryptography algorithm which is based on the factoring problem as well as Subset-Sum problem called a *Modified Subset-Sum over RSA public key cryptosystem (SSRPKC)*.

1.1 Euler's Phi-Function

Euler's phi-function, $\phi(n)$, which is sometimes called Euler's Totient Function; find the number of integer that are both smaller than n and relatively prime to n . This function $\phi(n)$ calculate the number of element in given set [3].

In this paper, we compare and evaluate the RSA cryptosystem and modified sub-set sum cryptosystem by implementing and running them on a computer. We investigate the issues of complexity, efficiency and reliability by running the algorithm with different sets of values. Moreover, comparisons will be done between these two algorithms given the same data as input.

The rest of paper is organized as follows- section 2, describes the RSA cryptosystem which depends on the factoring large integer numbers. In section 3, describe the introduction of sub-set sum cryptography which depends on the super-increasing order also called NP-complete problem. In section 4, we present our modified algorithm. In section 5, we compare both the algorithm- RSA and modified sub-set sum cryptosystem. A conclusion is shown in section 6.

II. RSA CRYPTOSYSTEM

RSA is based on the principle that some mathematical operations are easier to do in one direction but the inverse is very difficult without some additional information. In case of RSA, the idea is that it is relatively easy to multiply but much more difficult to factor. Multiplication can be computed in polynomial time where as factoring time can grow exponentially proportional to the size of the number. RSA consist of three steps [4]:-

Step 1) Key Generation Process

1. Generate two large random primes, p and q , of approximately equal size such that their product $n = p \times q$ is of the required bit length, e.g. 1024 bits.
2. Compute $n = p \times q$ and $\phi = (p-1) \times (q-1)$.
3. Choose an integer e , satisfying $1 < e < \phi$, such that $\gcd(e, \phi) = 1$.
4. Compute the secret exponent d , $1 < d < \phi$, such that $e \times d \equiv 1 \pmod{\phi}$.
5. The public key is (n, e) and the private key is (n, d) . Keep all the values d , p , q and ϕ secret.
6. n is known as the *modulus*.
7. e is known as the *public exponent* or *encryption exponent* or just the *exponent*.
8. d is known as the *secret exponent* or *decryption exponent*.

Public key (n, e) is published for every one and private key (p, q, d) must be kept secret. Then by using these keys encryption, decryption, digital signing and signature verification are performed.

Step 2) Encryption Process

Sender A does the following: -

1. Obtains the recipient B's public key (n, e) .
2. Represents the plaintext message as a positive integer m .
3. Computes the cipher text $c = m^e \pmod{n}$.
4. Sends the cipher text c to B.

Step 3) Decryption Process

Recipient B does the following: -

1. Uses private key (n, d) to compute $m = c^d \pmod{n}$.
2. Extracts the plaintext from the message representative m .

2.1 Security of RSA

The security of RSA cryptosystem is also broken by two attacks based on factorization attack and chosen-cipher text attacks [9].

A) Factorization Attack: The security of RSA is based on the idea that the modulus is so large that it is infeasible to factor it in a reasonable time. Bob select p and q and calculate $n=p*q$. Although n is

public, p and q are secret. If Eve can factor n and obtain p and q , she can calculate $\phi = (p-1) \times (q-1)$. Eve then calculate $d \equiv 1 \pmod{\phi}$ because e is public. The private exponent d is the trapdoors that Eve can use to decrypt any encrypted message [9].

B) Chosen-Cipher text Attack: A potential attack on RSA is based on the multiplicative property of RSA [9].

III. SUB-SET SUM CRYPTOGRAPHY

In computer science, the subset sum problem is an important problem in complexity theory and cryptography. The problem is this: given a set of integers, does the sum of some non-empty subset equal exactly zero. For example, given the set $\{-7, -3, -2, 5, 8\}$, the answer is yes because the subset $\{-3, -2, 5\}$ sums to zero. The problem is NP-Complete. There are two problems commonly known as the subset sum problem. The first is the problem of finding what subset of a list of integers has a given sum, which is an integer relation problem where the relation coefficients a_i are 0 or 1. The second subset sum problem is the problem of finding a set of n distinct positive real numbers with as large collection as possible of subsets with the same sum [4].

The subset sum problem is a good introduction to the NP-complete class of problems. There are two reasons for this [4]-

- It is a decision and not an optimization problem
- It has a very simple formal definition and problem statement.

IV. MODIFIED ALGORITHM

In this section we introduce a new approach for public key cryptosystem. Modified Subset Sum (MSS) is an asymmetric-key cryptosystem in which two keys are required: a public key and a private key. Furthermore, unlike RSA, it is one-way, the public key is used only for encryption, and the private key is used only for decryption. Thus it is useless for authentication by cryptographic signing. Modified algorithm consist of three steps-

Step 1) Key Generation Process

1. Generate two large random primes, p and q , of approximately equal size such that their product $m = p \times q$ is of the required bit length, e.g. 1024 bits. (From Big Integer library function of Java)
2. Compute $m = p \times q$ and $\phi = (p-1) \times (q-1)$.
3. Choose an integer e , satisfying $1 < e < \phi$, such that $\gcd(e, \phi) = 1$.
4. Compute the secret exponent d , $1 < d < \phi$, such that $e \times d \equiv 1 \pmod{\phi}$.
5. Choose a superincreasing set $A = (a_1, \dots, a_n)$
6. Choose an integer M with $M > \text{SUM}_{i=1 \dots n}(a_i)$. M is called the modulus.
7. Choose a multiplier W such that $\gcd(M, W) = 1$ and $1 \leq W < M$. This choice of W guarantees an inverse element U : $UW = 1 \pmod{M}$
8. To get the components b_i of the public key B , perform $b_i = a_i * W \pmod{M}$, $i = 1 \dots n$

The superincreasing property of A is concealed by *modular multiplication*.

The public key is (B, n, e) and the private key is (A, M, W, n, d) . Keep all the values d , p , q and ϕ secret.

Public key is published for every one and private key must be kept secret. Then by using these keys encryption and decryption are performed.

Step 2) Encryption of Message

Sender A does the following: -

1. The length of a message to be encrypted is fixed by the parameter n prior to encryption; a possibly larger message p has to be divided into n -bit groups.
2. Let $p = (p_1, p_2, \dots, p_n)$ the message to be encrypted.
 - The ciphertext c is obtained by computing $c = b_1 p_1 + b_2 p_2 + \dots + b_n p_n$
 - Computes the cipher text $c_1 = c^e \pmod{n}$.
 - Sends the cipher text c_1 to B.

Step 3) Decryption of Message

Recipient B does the following: -

1. Uses private key and first compute $m_1 = c_1^d \bmod n$.
2. First compute $c' = Um_1 \bmod M = W^{-1}c \bmod M$
3. Now solve (A, c') . Because A is superincreasing, (A, c') is easily solvable. Let $X = (x_1, \dots, x_n)$ be the resulting vector and $p_i = x_i$ and $p = (p_1, \dots, p_n)$ is the plaintext.

V. COMPARISON OF RSA AND SSRPKC CRYPTOSYSTEMS

5.1 Simulation Results

For the simulation purpose, SSRPKC cryptosystem is implemented as a user-friendly GUI. This GUI application is implemented using Java Big Integer library functions [5]. In this application, one can either enter the prime numbers or can specify the bit length of the prime numbers to generate automatically. Big Integer library provides operations for modular arithmetic, GCD calculation, primarily testing, prime generation, bit manipulation, and a few other miscellaneous operations. The simulation of the algorithm, implemented in JAVA [5], running on a 2 GHz P-IV Processor and 512 MB RAM, using a 1000 characters long message for encryption/decryption. The PKC algorithms (RSA & SSRPKC) have some important parameters affecting its level of security and speed [6]. The complexity of decomposing the modules into its factors is increased by increasing the module length. This also increases the length of private key and hence difficulty to detect the key. Another parameter is the number of items in set A . As the number of items in set A increases, the size of the message which is encrypted at a time also increases, hence the security also increases as well as difficulty of detecting the private set A from public set B also increases. The RSA and SSRPKC parameters are changed one parameter at a time and the others are kept fixed to study the relative importance. The key generation, encryption, decryption time is depends on the speed of the processor and the RAM. Table 5.1 shows the simulation results of both the algorithms.

5.1.1 Changing the modulus length:

Changing the modulus affects the other parameters of the algorithms as shown in Table 5.1. It is clear here that increasing the modulus length (bits) increases the bit length of their factors and so the difficulty of factoring them into their prime factors. Moreover, the length of the secret key (d) increases at the same rate as n -bit increases. As a result, increasing the n -bit length provides more security. On other hand by increasing the n -bit length increases the values of key generation time, encryption time and decryption time. Hence increasing the n -bit length increases the security but decreases the speed of encryption, decryption and key generation process as illustrated by Figure 5.1 and 5.2.

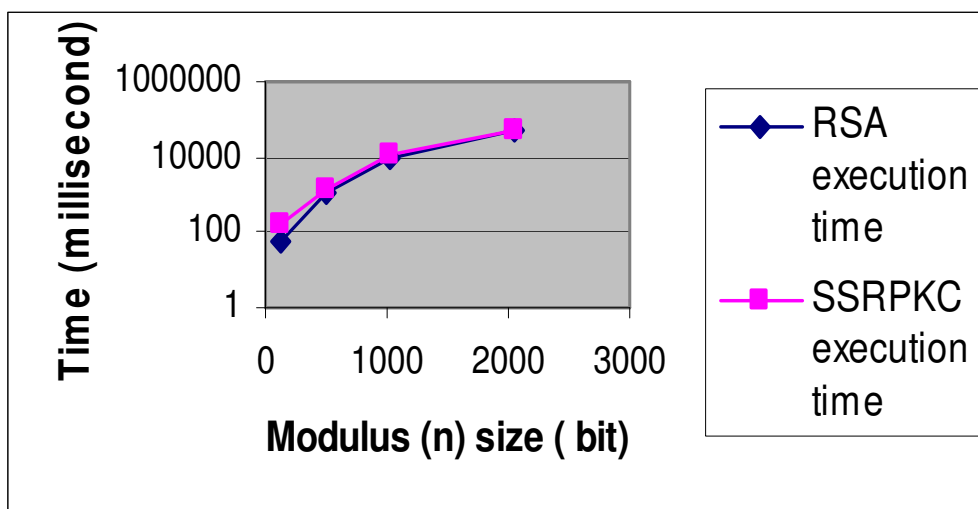
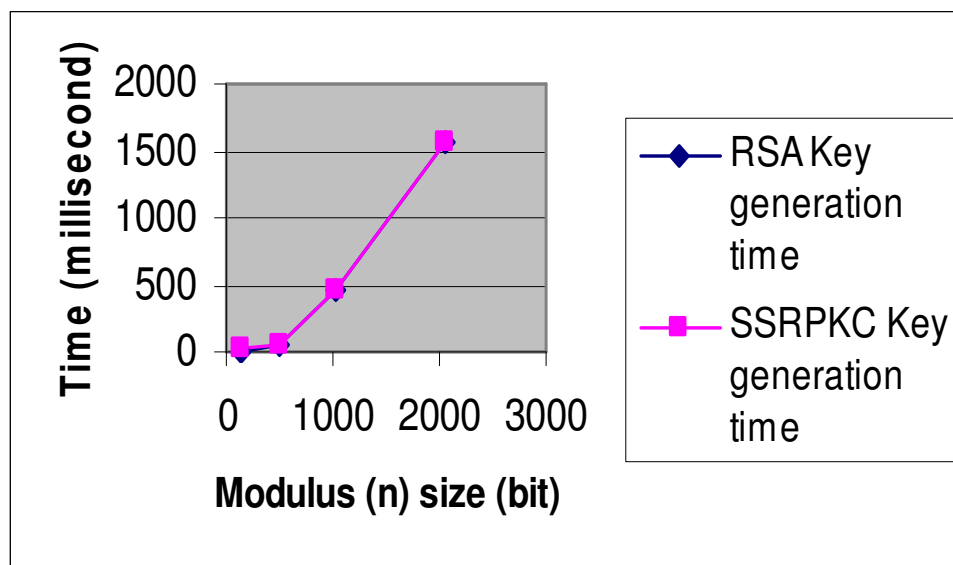


Figure1. Modulus size v/s RSA & SSRPKC algorithm's execution time, taking number of items in set A 128 and size of Public Key 128 bit.

Table 1: Effect of changing the modulus length and size of set A on the size of Private key, Key generation time, Encryption time and Decryption time, while the size of Public key has kept constant (128 bit)

Size of N (bit)	Size of d (bit)	Number of element in set A	SSRPKC				RSA			
			key generation time (ms)	Encryption time (ms)	Decryption time (ms)	Total Execution time(ms)	Key Generation on time (ms)	Encryption time (ms)	Decryption time (ms)	Total Execution time(ms)
128	128	32	16	235	78	329	16	94	62	172
128	128	64	15	94	47	156	0	16	47	63
512	512	32	125	1203	1766	3094	109	563	1719	2391
512	512	64	63	344	875	1282	63	141	859	1063
512	512	128	78	172	453	703	47	78	422	547
1024	1024	32	688	5407	11328	17423	688	1719	12172	14579
1024	1024	64	453	6593	5735	12781	453	2968	5688	9109
1024	1024	128	562	516	2859	3937	515	219	3344	4078
1024	1024	512	12812	187	781	13780	281	47	735	1063
2048	2048	32	3735	9563	85140	98438	3719	3688	85672	93079
2048	2048	64	1563	3625	42437	47625	1563	1688	43234	46485
2048	2048	128	7125	6266	20734	34125	7078	3829	21406	32313
2048	2048	512	17797	797	5281	23875	7703	375	6172	14250
2048	2048	1024	29704	422	2797	32923	2891	203	3406	6500

**Figure 2.** Modulus size v/s Key generation time, taking size of Public key 128 bit and number of items in set A are 128

5.1.2 Changing the number of items in set A:

On the basis of simulation results of Table 5.1, following Figure 5.3 shows the effect of number of items on encryption and decryption time of both the algorithms. Here key generation time of SSRPKC algorithm depends on the number of items in set A and as the number of items increases key generation time also increases that's why for more than 500 items in set A, execution time of SSRPKC algorithm also increases. RSA algorithm's execution time doesn't depend on set-A items.

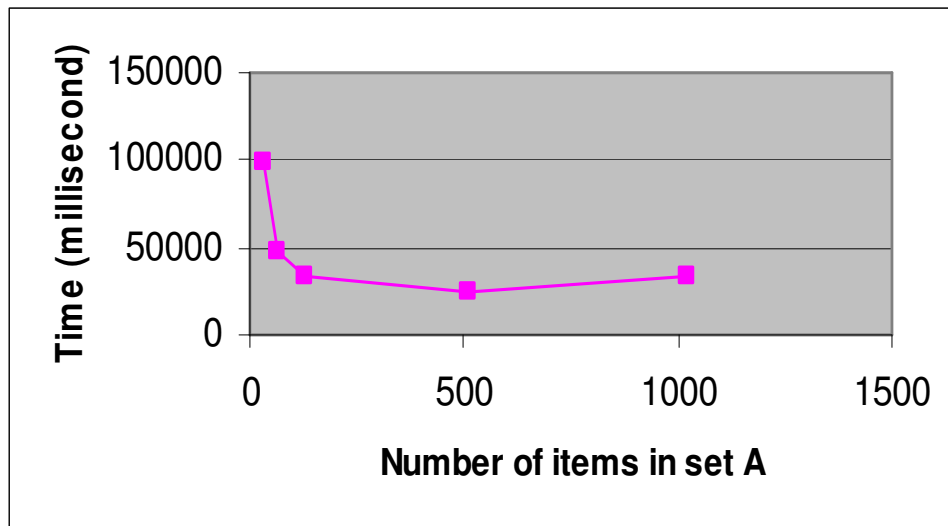


Figure 3. Number of items in set A v/s SSRPKC algorithm's execution time, taking size of modulus 2048 bit and size of private 128 bit.

5.2 Complexity analysis of SSRPKC cryptosystem

The RSA cryptosystem is well-known asymmetric cryptosystem in which the computational complexity for both of the encryption and the decryption are of the order of k^3 , where k is the number of bits of the modulus n [7]. The computational complexity of subset sum problem is $O(k)$, if a super-increasing set with greedy approach is used where k is the number of items in set-A [8]. So in this way, computational complexity of SSRPKC cryptosystem is on the order of $O(k + k^3)$ i.e. $O(k^3)$ or encryption and on the order of k^3 for decryption. So the computational complexity of SSRPKC is equivalent to RSA cryptosystem. The simulation results of both the algorithms shows that the execution time of SSRPKC is 1.2 times more than RSA.

5.3 Security analysis of SSRPKC cryptosystem

Two possible approaches to attacking the SSRPKC / RSA algorithms are as follows [1, 9]:

1. *Brute force*: This involves trying all possible private keys.
2. *Mathematical attacks*: These are based on factoring the product of large primes such as factor n into its prime factors p and q respectively then calculating ϕ , which, in turn, enables determination of $d = e^{-1} \bmod \phi$.

Estimated resources needed to factor a number within one year are as follows [9].

Table 2: Recourses needed to factor a number within one year

Size of number (Bits)	PCs	Memory
430	1	128 MB
760	215,000	4 GB
1020	342×10^6	170 GB
1620	1.6×10^{15}	120 TB

As the SSRPKC cryptosystem is based on the hybrid combination of the subset sum problem and RSA cryptosystem, so one have to factor the modulus into its primes as well as find the secret set-A to break the SSRPKC algorithm [2, 7]. If RSA which is based on single modulus, is broken in time x and Subset sum based algorithms is broken in time y then the time required to break SSRPKC algorithm is $x*y$. So the security of SSRPKC algorithm is increased as compare to RSA algorithm and it shows that the SSRPKC algorithm is more secure for *Mathematical attacks* [1].

As in SSRPKC double decryption is performed and unlike RSA that is not only based on private key but also based on the subset sum problem so one can't break SSRPKC only guessing the private key only. So it shows that SSRPKC algorithm is more secure as compare to RSA for *Brute force attack* [1]. The RSA cryptosystem is based on the Integer Factorization which is break by the Pollard rho method [4], Pollard $p-1$ method [4] and Fermat method [4]. Because of these methods the security of RSA cryptosystem is broken by the Factorization attack [3] and Chosen Cipher Text attack [3].

As the SSRPKC cryptosystem uses dual modulus, so for breaking this one have to factor both the modulus. That's why our cryptosystem provides the far better security against the factorization methods and attacks on RSA cryptosystem.

VI. CONCLUSION

This paper presents an extended version of Subset-Sum problem over RSA algorithm called Extension of cryptosystem through subset-sum over RSA. It relies on facts of mathematics that indicates that given a very large number it is quite impossible in today's aspect to find two prime numbers whose product is the given number. The size of number increases, the possibility for factoring the number decreases. In RSA, if one can factor modulus into its prime numbers then he can get the private key too. To improve the security, SSRPKC cryptosystem is developed. The disadvantage of new cryptosystem is that, unlike RSA it can not be used for authentication as it is based on the one way function. Another disadvantage is the slow down of execution process as compare to RSA. But it is clear from the simulation results that it is more secure than the RSA algorithm and our cryptosystem provides the far better security against the factorization methods and attacks on RSA cryptosystem.

REFERENCES

- [1] William Stallings, "Cryptography and Network Security", ISBN 81-7758-011-6, Pearson Education, Third Edition, pages 42-62, 121-144, 253-297.
- [2] Ralph C. Merkle, Martin E. Hellman. "Hiding Information and Signatures in Trapdoor Knapsacks", IEEE Transactions on Information Theory, vol. IT-24, 1978, pp. 525-530.
- [3] Behrouz A. Forouzan, Debdeep Mukhopadhyay, "Cryptography and Network Security", 2nd Edition, TMH.
- [4] Alfred J Menezes, Paul C van Oorschot, Scot A Vanstone, Handbook of Applied Cryptography, CRC Press, 1997.
- [5] Neal R. Wagner, "The Laws of Cryptography with Java Code", Technical Report, 2003, pages 78-112.
- [6] Allam Mousa, "Sensitivity of Changing the RSA Parameters on the Complexity and Performance of the Algorithm", ISSN 1607 – 8926, Journal of Applied Science, Asian Network for Scientific Information, pages 60-63, 2005.
- [7] RSA Laboratory (2009), "RSA algorithm time complexity", Retrieved from <http://www.rsa.com/rsalabs/node.asp?id=2215> (4 April 2011).
- [8] Adi Shamir, A Polynomial Time Algorithm for Breaking the Basic Merkle-Hellman Cryptosystem. CRYPTO 1982, pp279–288.
- [9] CHUK, Chinese university (2009), "RSA Algorithm security and Complexity", Retrieved from [http://www.cse.cuhk.edu.hk/~phwl/mt/public/archives/old/ceg50 10/rsa.pdf](http://www.cse.cuhk.edu.hk/~phwl/mt/public/archives/old/ceg50%2010/rsa.pdf) (26 Jan. 2011)
- [10] Wenbo Mano, "Modern Cryptography Theory and Practice," Prentice Hall.
- [11] Bryan Poe, "Factoring the RSA Algorithm", Mat / CSC 494, April 27, 2005, pages 1-6.
- [12] Adi Shamir, A Polynomial Time Algorithm for Breaking the Basic Merkle-Hellman Cryptosystem. CRYPTO 1982, pp279–288
- [13] A Method for Obtaining Digital Signatures and Public-Key Crypto Systems by R.L. Rivest, A. Shamir, and L. Adleman.
- [14] Menezes, A. J., Van Oorshot, and Vanstone, P.C. S.A. Handbook of Applied Cryptography, CRC press, 1997.
- [15] B. Schneizer, Applied Cryptography, New York: John Wiley, 1994.

- [16] <http://www.certicom.com>, "The Elliptic Curve Cryptosystem," September 1997, (dated 02-04-2010)
[17] J H Moore, Protocol failures in Cryptosystems, Contemporary Cryptology, The science of Information Integrity, Ed. G J Simmons, 541-558, IEEE Press 1992.
[18] Diffie, M Hellman, "New Directions in Cryptography", IEEE Transactions on Information Theory, Vol 22, 1976.

Author's Biography

Sonal Sharma was born on 23 May 1984. She is M.Tech. Student in SBCET, Jaipur (Rajasthan). She has completed B.E. (I.T.) in 2006 from University of Rajasthan, Jaipur.



Saroj Hiranwal was born on 20 Jan 1982. She has done B.E.(I.T.) in 2004 and MTECH(CSE) in 2006. Her teaching Experience is 6.5 year in the organization - Sri Balaji College of Engineering and Technology, Jaipur with the designation of Reader and HEAD.

Prashant Sharma was born on 04 June 1985. He is the M.Tech. Student in SBCET, Jaipur (Rajasthan). He has completed B.E.(I.T.) in 2007 from University of Rajasthan, Jaipur.



A COMPACT DUAL BAND PLANAR RMSA FOR WLAN/WIMAX APPLICATIONS

C. R. Byrareddy¹, N. C. Easwar Reddy², C. S. Sridhar³

¹ Assistant Prof., Dept of E&C. Engg., Bangalore Institute of Technology, Bangalore, India

² Prof., Dept of E&C. Engg., S.V. U. College of Engg., Tirupathi, India

³ Professor, Dept of E&C. Engg., Bangalore Institute of Technology, Bangalore, India

ABSTRACT

Presentation of a compact dual band planar rectangular microstrip antenna (RMSA) antenna for a WLAN (2.4GHz IEEE standards 802.11b/g)/WiMAX(2.6GHz IEEE standards 802.16e) applications. The two resonant modes of the presented RMSA antenna are associated with various length and width of the planar strips in which a centre strip contributes for the lower resonant frequency 2.4GHz(2.26-2.4GHz with impedance bandwidth 240MHz) and two lateral strips contributes for the higher resonant frequency 2.8GHz(2.73-2.95GHz with impedance bandwidth 220MHz). By proper adjustment of the coupling between the two lateral strips and embedded centre strip enables the operation of dual band with a -10dB return loss, a near directive radiation pattern and a good antenna gain with sufficient bandwidth. The antenna is simulated using Ansoft HFSS and fabricated on an FR4 substrate with dielectric constant 4.4 and thickness 1.6mm occupies an area of 65mm x 50mm. The simulation results are found to be in good agreement with the measured results. The proposed antenna is suitable for wireless communication applications requiring a small antenna.

KEYWORDS: Rectangular Microstrip Antenna (RMSA), Wireless Local Area Network (WLAN), WiMAX, Strips, monopole dual band.

I. INTRODUCTION

Rapid progress in wireless communication services have led to an enormous challenge in antenna design. Patch antennas for dual and multi frequency band operation has increasingly become common, mainly because of many advantages such as low profile, light weight, reduced volume and compatibility with microwave integrated circuits (MIC) and monolithic microwave integrated circuit (MMIC). WLAN is one of the most important applications of the wireless communication technology that takes advantage of licence free frequency bands [ISM] due to high speed connectivity between PCs, laptops, cell phones and other equipments in environments. In the near future WiMAX technology with different standards is going to occupy the market. Wireless data services have evolved and continues to grow using various technologies, such as 2G/3G. The impact of such diverse technologies is on the use of frequency band in different technologies will need to occupy different frequency allocations, Such as WLAN/WiMAX, it likely to be prominent candidate to serve for wireless data in near future. Therefore there is a need to develop a dual band antenna for both WLAN/WiMAX applications occupying 2.4/2.6GHz frequency bands.

Above several papers on dual band antennas for IEEE standards have been reported. [1-2] proposed printed double T-monopole antenna can cover the 2.4/5.2 GHz WLAN bands and offers narrow band width characteristics and planar branched monopole antenna for DCS/2.4GHz. For WLAN it can provide excellent wide frequency band with moderate gain. [3] The proposed planar monopole antenna is capable of generating good Omni directional monopole with radiation in all the frequency bands. [4-5] proposed printed dipole antenna with parasitic element and Omni-directional planar

antenna for WiMAX applications, can operate either in wide band or dual band, which cover 3.25-3.85, 3.3-3.8 and 5.15-5.85 GHz with return loss of -10dB.[6] Broad band printed L-shaped antenna for wireless communication is reported with good radiation patterns and better return loss.[7] physical design features proper geometry and dimension for microstrip antenna array using transformer $\lambda/4$ for the feed line matching technique.[8] proposes compact terminal antenna incorporates open-end slots in the ground plane, which reduces size and operates at acceptable band width.[10] use of various feeding techniques can give dual or multiband operations.

In this paper a compact dual band antenna structure for WLAN and WiMAX are proposed. The proposed antenna is simple to design and offer an effective control over two operating bands by controlling the dimensions of three rectangular strips. The antenna can easily be fed using a 50Ω probe feed with transformer $\lambda/4$ technique for impedance matching. Also the planar RMSA structure antenna is attractive from the package point of view. The advantage of $\lambda/4$ technique feeding method is to match the transmission line characteristics impedance to the input impedance.

II. ANTENNA GEOMETRY AND DESIGN

The geometry of the proposed antenna structure is shown in figure 1. It is etched on a substrate of dielectric constant $\epsilon_r=4.4$ and thickness $h=1.6\text{mm}$ with tangent loss 0.09. The antenna has ground plane dimension of length $L_g=50\text{mm}$ and width $W_g=65\text{mm}$. The radiating structure consists of three rectangular strips with dimensions of length $l_p=28.5\text{mm}$ and centre strip width $w_{p1}=18\text{mm}$, lateral strips width $w_{p2}=10\text{mm}$ with a slot gap width $g=0.5\text{mm}$. The centre strip is fed by a designed 50Ω microstrip line width 0.5mm , the optimum feed point antenna is $\lambda/4$ Transformer method with 3mm width and 0.2mm height for good impedance matching. Thus it can be connected with a SMA connector. The resulting antenna resonates at 2.4GHz and 2.8GHz . From simulation and experimental studies, it is found that the dimensions of the middle rectangular strip are optimized to resonate at 2.4GHz while dimensions of the lateral symmetrical strips are optimized to resonate at 2.8GHz . Thus the proposed antenna provides effective control of the two operating bands. In addition, ground plane dimensions are also optimized to achieve the desired dual band operation, as it affects the resonant frequencies and operating band widths in two bands.

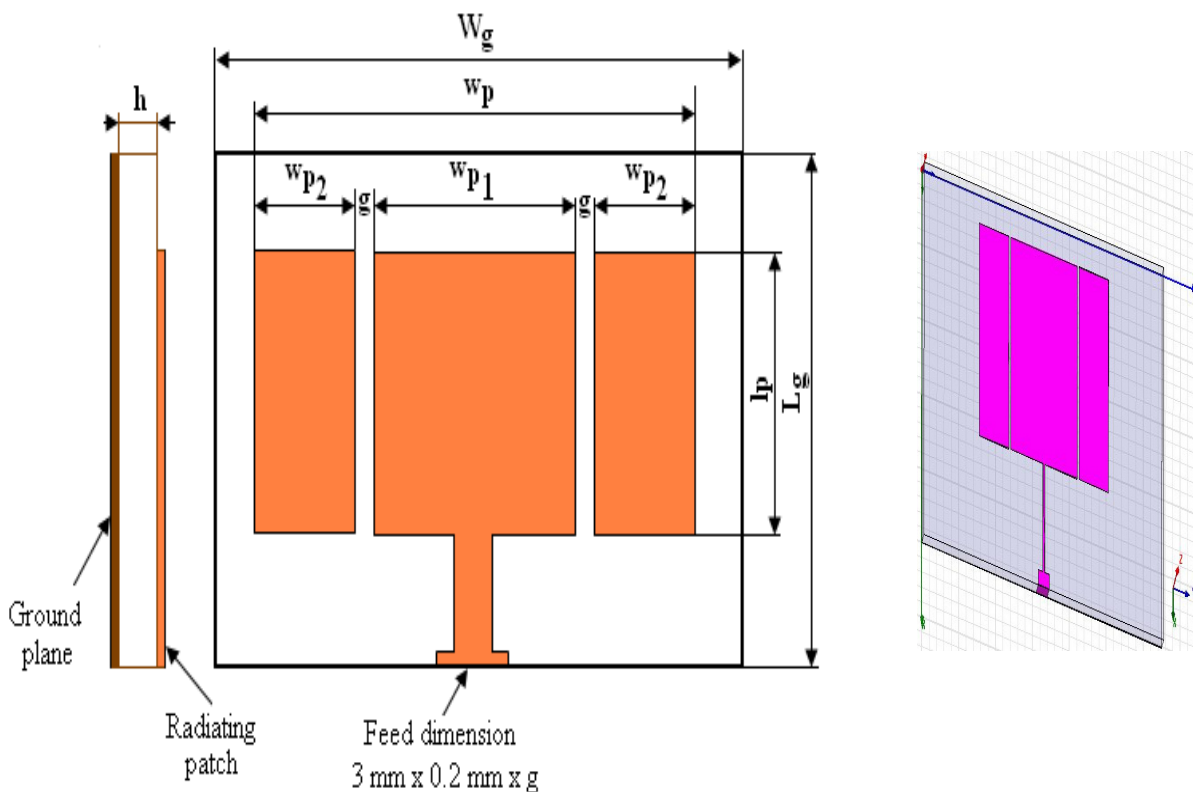


Figure.1. Geometry of the proposed planar RMSA, a) side view, b) top view, c) simulated view.

III. SIMULATION AND EXPERIMENTAL RESULTS

Ansoft HFSS is used for simulation and analysis of the structure. Simulation process involves setting up of the optimum geometric dimension to satisfy the desired centre frequency, as well as the bandwidth requirement for specified return loss requirements in each band. The simulation was iteratively conducted until the desired results were found. The proposed antenna is fabricated by using photolithographic process which gives goods accuracy for the etched patterns. The fabricated antenna parameters are measured by experimental set up characterization which consists of vector network analyzer from Hewlet Packard, with S-parameter test-set and an anechoic chamber. Radiation patterns, E-field &H-field, S-parameters and gain were measured. The following sections describe the details of measured and simulated results. Measurement of return loss is most important because our main interest in this research is to produce dual band characteristics within the specified centre frequency with sufficient impedance bandwidth requirement.

3.1 Return loss

Figure.2 shows the simulated and experimental return loss of the proposed dual band antenna. From the simulation, the impedance bandwidth of the lower frequency band determined by -10dB return loss approximately 240MHz of bandwidth (2.26-2.50GHz), which is 13% of bandwidth for the frequency band of 2.4 GHz. For the upper frequency band the impedance bandwidth is approximately 220MHz (2.73-2.95GHz), i.e. about 10% for the frequency band of 2.8GHz. The centre frequencies the two bands are determined by adjusting the rectangular dimensions of strips. To achieve the maximum results the gap distance between the two strips adjusted and length of the microstrip feed line 50 ohm need to be controlled. The experimental curve shows that a dual band is obtained for both the resonance with good matching.

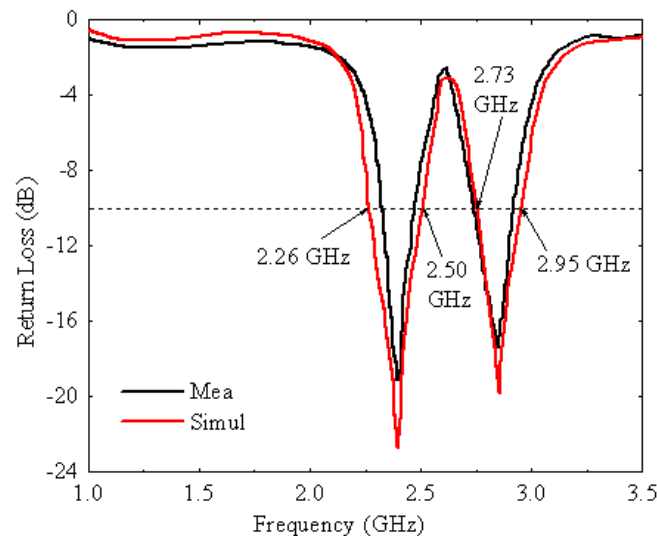


Figure.2. Simulated and measured return loss characteristics of antenna structure.

3.2 Radiation Patterns

The simulated and measured radiation patterns of the proposed antenna operating at 2.4GHz and 2.8GHz are shown in figure 3 & 4 respectively. It is found that the antenna has relatively stable radiation patterns over its operating band, a near Omni directional pattern is obtained in the two bands. Because of symmetry in the antenna structure the radiation patterns are as good as those of a conventional monopole. In the elevation plane (azimuth angle) as shown in the plots, asymmetrical radiation patterns are seen in the x-y and x-z planes. The measured radiation patterns are stable and quasi -Omni directional in the entire operational band which is highly suitable for the proposed modern wireless communication bands.

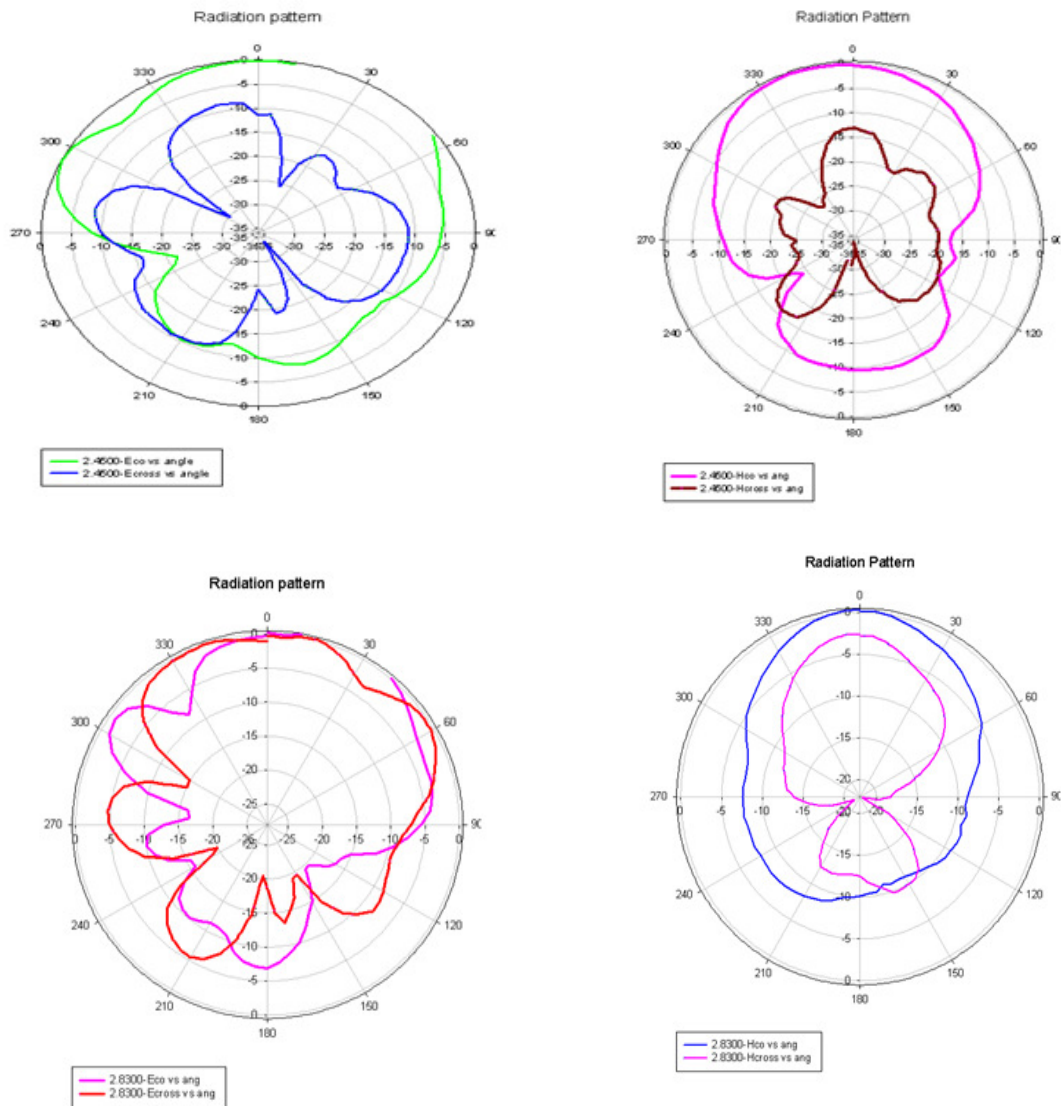


Figure.3. Simulated co-polar and cross-polar radiation patterns of the planar RMSA AT 2.4GHz and 2.8GHz.

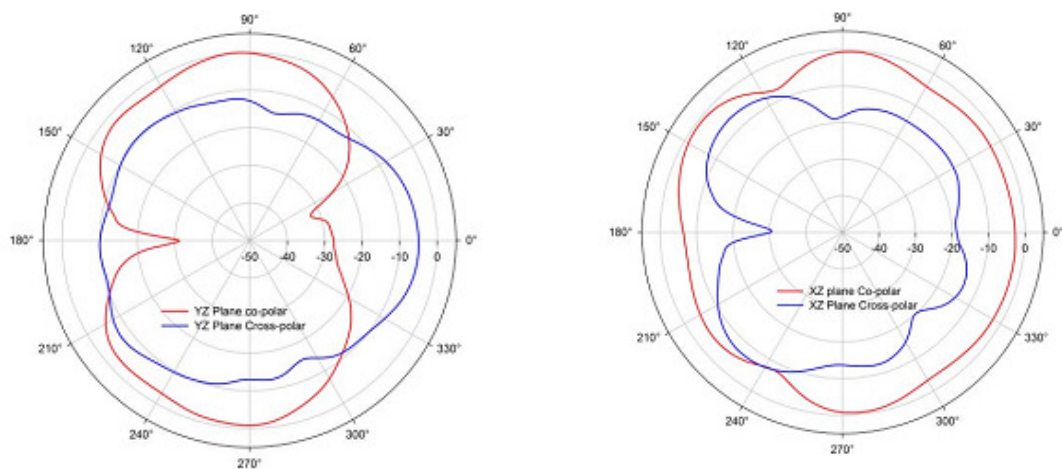


Figure.4. Measured radiation patterns at 2.4GHz and 2.8GHz.

3.3 Current Distribution Characteristics and 3-D plot

A better understanding of the antenna behaviour can be obtained by analyzing the current distributions at peak resonant frequency 2.4 GHz, as shown in figure 5. It is evident from that at 2.4GHz the central strip acts as a quarter wave monopole where as for the higher resonance the predominant effect is seen at the edges of the lateral strips.

Figure.6 shows the 3D simulated radiation pattern at 2.4GHz. It is found that the planar antenna provides almost Omni directional radiation coverage and can be used for WLAN applications.

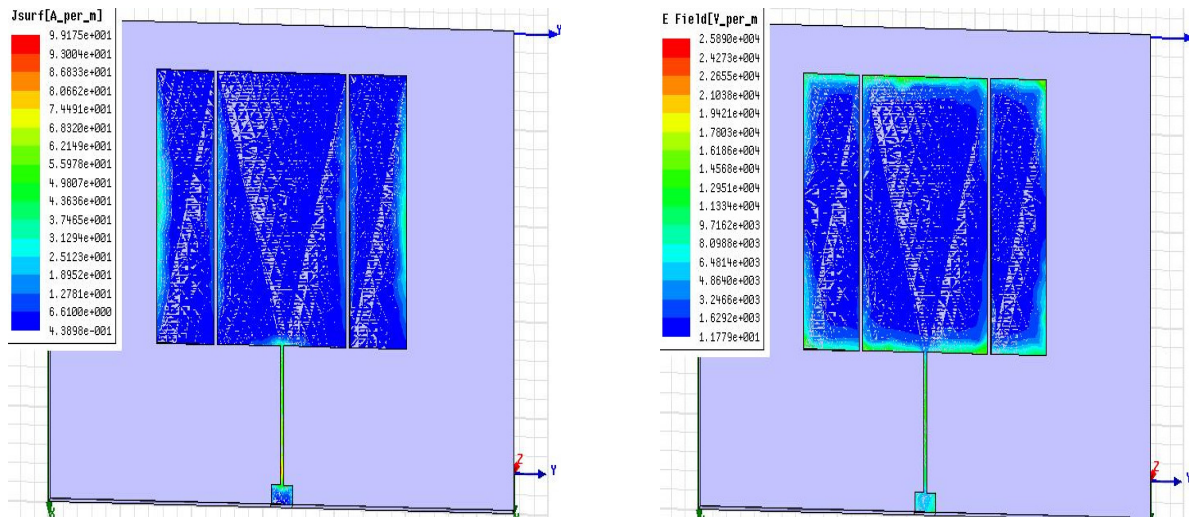


Figure.5 current distribution and e-field distribution of the planar RMSA

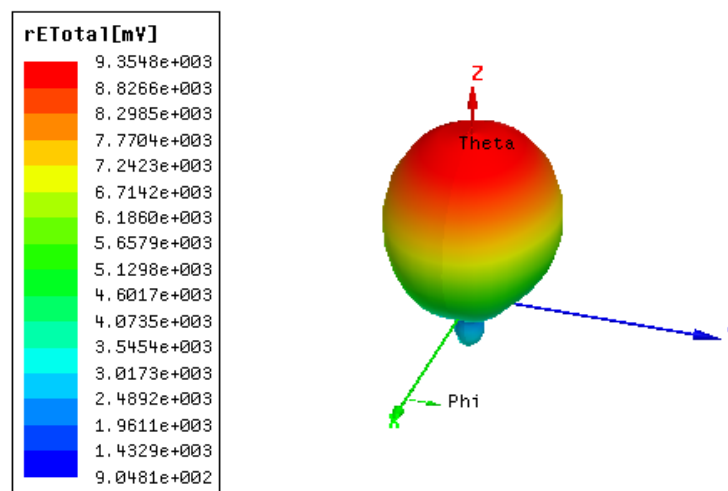


Figure.6 Simulated 3D radiation pattern of the planar RMSA

3.4 Gain characteristics

The peak gain of the antenna measured at each frequency points by comparison method. Figure.7 shows the measured antenna gain versus frequency. Average gain at different frequencies of both bands are shown where as at 2.4GHz frequency band is approximately 2.95dBi and at 2.8 GHz frequency band is approximately 3.8dBi and then continue for higher frequencies. Figure.8 shows photo of the fabricated planar RMSA and tested by measuring its parameters particularly return loss and antenna gain, to validate the simulation result as well as to verify the antenna design specification.

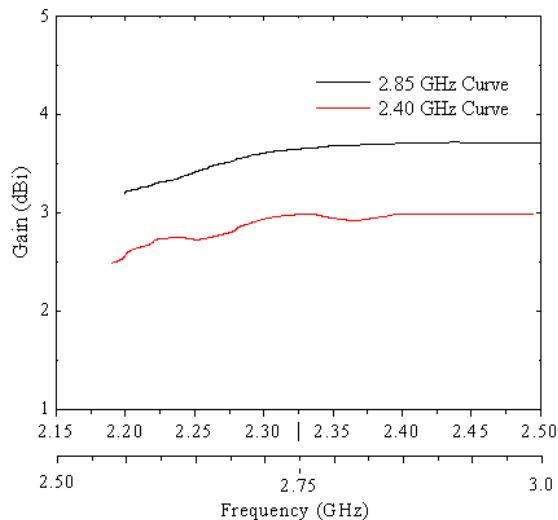


Figure.7 Measured peak antenna gain for the dual band planar RMSA at 2.4GHz and 2.8GHz.

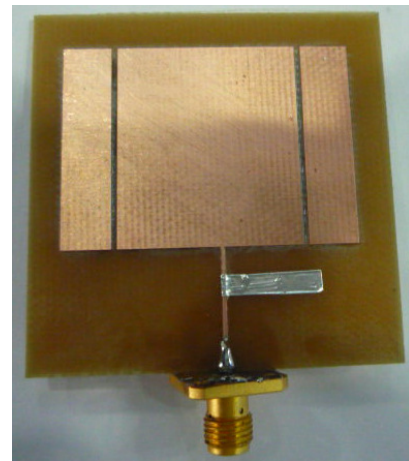


Figure.8 Photo of fabricated planar RMSA

IV. CONCLUSION

A study of construction and experimental verification of compact dual band planar RMSA structure for the operation in the 2.4/2.8GHz WLAN/WiMAX bands is presented. The dual band operation is achieved by using planar three rectangular strips. A good band width characteristics is observed in each strips which covers 2.26- 2.5GHz and 2.73-2.95GHz bands with an impedance bandwidth of nearly 240MHz at lower band and at the upper band the impedance bandwidth 220MHz when compared with the simulation , measured results exhibits good agreement, except for small variation due to measurement error. From our design it is found that the lower resonant band is due to centre strip and the higher resonant band due to lateral strips. This is because good impedance matching, gain and radiation pattern can be obtained by tuning the coupling between the rectangular strips and the gap, as well as the feed length of the feed line, Hence the proposed antenna is suitable for wireless local area network (WLAN) and multichannel multipoint distribution service (MMDS)WiMAX communication applications.

ACKNOWLEDGEMENT

The authors are thankful to CUSAT, Cochin for providing lab facilities to fabricate and test.

REFERENCES

- [1] Y.L.Kuo and K.L.Wong, "Printed double -T monopole antenna for 2.4/5.2 GHz dual band WLAN operations", IEEE Transaction antenna and propagation, vol.51,no 9,pp.2187-2192,sept.2003.
- [2] Suma,M.N., Raj.R.K, Joseph.M, Bybi.P.C, Mohanan.p, "A Compact dual band planar branched monopole antenna for DCS/2.4 GHz WLAN applications". Microwave and wire less components letters, IEEE,vol.16,issue.5,pp 275-277,2006.
- [3] Das.S.K, Gupta.R.K, Kumar.G, "Dual band planar monopole antenna ", IEEE antenna and propagation international symposium 2006,pp.1705-1708, July 2006.
- [4] Chen,W.-S., Yu, Y.-H. "Dual band printed dipole antenna with parasitic element for WiMAX applications". Electronics letters vol.44, issue. 23, pp.1338-1339 ,IET journals, 2008.
- [5] Ting-Ming Hsueh,Heng-Tung Hsu, His-TsengChou, Kwo-Lun Hung, "Dual band Omni- directional planar antenna for WiMAX applications". IEEE Antenna and propagation society international symposium, 2008, PP.NO. 1-4, AP-S 2008,
- [6] Qinjiang Rao and Tayeb A. Denidi, "New broad band Dual printed inverted L-shaped monopole antenna for Tri-band wireless applications", microwave and Optical technology letters, vol.49, no.2, pp.278-280 Feb.2007
- [7] Surjati, I et al, "Dual Band Triangular Microstrip Antenna Using Slot Feed By Electromagnetic Coupling", Proceeding Quality In Research, Faculty Of Engineering, University Of Indonesia,2009 .

- [8] R.Hossa, .Byndas, and M.E.Bialkowski, "Improvement of compact terminal antenna performance by incorporating open-end slots in ground plane", IEE E Microwave and wireless components letters, vol.14,no.6,pp.283-285,2004.
- [9] Pozar, David M., Microwave Engineering, Second Edition, Wiley, New York 1998.
- [10] S. Rhee. & G. Yun "CPW fed slot antenna for triple-frequency band operation", Electronics Letters, Vol. 42, No.17, pp. 952-953, 2006.
- [11] HFSS V-12 tool used, HFSS user manual, Ansoft Corporation, USA.

Author Profiles

C. R. BYRA REDDY born in Karnataka, India in 1967. He graduated from Bangalore University with B.E. degree in Instrumentation technology and the M.E. degree in Electronics in 1990 and 1999 respectively. He is currently Assistant professor in department of Electronics & Communication engineering, Bangalore institute of technology. He is a Ph.D. candidate at SV University of engineering college, Tirupathi. Research interest area is Microwave communication, Antennas and Wireless communication, with a keen interest Includes Analysis and design of patch antenna for Wireless communication. He has published 5 papers in national/international and in journals. He has presented paper at IconSpace2011 in Malaysia.



N. C. ESWAR REDDY born in Andhra Pradesh, India in 1950. He received the B.E degree from Andra University in 1972. He did his M.Tech. Degree from IIT Delhi in 1976. He did his Ph.D. from S.V. University in 1985. He joined S.V. College of Engineering as a lecturer in 1976, he has served as reader, professor and principal in the same college. His area of interest Microwave Engineering, Microprocessor, Bio signal processing and antennas. He has guided three Ph.D. candidates. He has published more than 32 papers both national and international journals, and has attended more than 20 international conference. He is Member of ISTE, IETE and Expert member of AICTE, Selection committee member of selection committee for all the university in and around Andhra Pradesh.



C. S. SRIDHAR born in Bellary, Karnataka and graduated with B.E degree from Andhra university in 1962, M.Sc Engineering from Madras University in 1966 and PhD from IIT Madras in 1975. He has been teaching since 1962 and has research interest in microwave antennas, signal processing architecture and VLSI Designs. He has published more than 35 papers national and international journals; He attended more than 60 international conferences. He is a life member of IETE and a member of IEEE.



VLSI ARCHITECTURE FOR LOW POWER VARIABLE LENGTH ENCODING AND DECODING FOR IMAGE PROCESSING APPLICATIONS

Vijaya Prakash. A.M¹ & K.S. Gurumurthy²

¹ Research Scholar, Dr. MGR University, Chennai, Faculty Department of ECE, BIT, Bangalore, India.

² Professor, Department of ECE, UVCE, Bangalore, India.

ABSTRACT

The image data compression has been an active research area for image processing over the last decade [1] and has been used in a variety of applications. This paper investigates the implementation of Low Power VLSI architecture for image compression, which uses Variable Length Coding method to compress JPEG signals [1]. The architecture is proposed for the quantized DCT output [5]. The proposed architecture consists of three optimized blocks, viz, Zigzag scanning, Run-length coding and Huffman coding [17]. In the proposed architecture, Zigzag scanner uses two RAM memories in parallel to make the scanning faster. The Run-length coder in the architecture, counts the number of intermediate zeros in between the successive non-zero DCT coefficients unlike the traditional run-length coder which counts the repeating string of coefficients to compress data [20]. The complexity of the Huffman coder is reduced by making use of a lookup table formed by arranging the {run, value} combinations in the order of decreasing probabilities with associated variable length codes [14]. The VLSI architecture of the design is implemented [12] using Verilog HDL with Low Power approaches. The proposed hardware architecture for image compression was synthesized using RTL complier and it was mapped using 90nm standard cells. The Simulation is done using Modelsim. The synthesis is done using RTL compiler from CADENCE. The back end design like Layout is done using IC Compiler. Power consumptions of variable length encoder and decoder are limited to 0.798mW and 0.884mW with minimum area. The Experimental results confirms that 53% power saving is achieved in the dynamic power of Huffman decoding [6] by including the lookup table approach and also a 27% of power saving is achieved in the RL-Huffman encoder [8].

KEYWORDS: Variable Length Encoding (VLE), Variable Length Decoding (VLD), Joint Photographic Expert Group (JPEG), Image Compression, Low Power Design, Very Large Scale Integration (VLSI).

I. INTRODUCTION

Image data compression refers [4] to a process in which the amount of data used to represent image is reduced to meet a bit rate requirement (below or at most equal to the maximum available bit rate) while the quality of the reconstructed image satisfies a requirement for a certain application and the complexity of computation involved is affordable for the application [18]. The image compression can improve the performance of the digital systems by reducing the cost of image [22] storage and the transmission time of an image on a bandwidth limited channel, without significant reduction in the image quality [15].

This paper investigates the implementation of Low Power VLSI architecture for image compression [8] which uses variable length coding method for image data compression, which could be then used for practical image coding systems to compress JPEG signals [1].

Variable length coding [2] that maps input source data on to code words with variable length is an efficient method to minimize average code length. Compression is achieved by assigning short code words to input symbols of high probability and long code words to those of low probability [14]. Variable length coding can be successfully used to relax the bit-rate requirements [21] and storage spaces for many multimedia compression systems. For example, a variable length coder (VLC) employed in MPEG-2 along with the discrete cosine transform (DCT) results [16] in very good compression efficiency.

Since early studies have focused only on high throughput variable length coders [6], low-power variable length coders have not received much attention. This trend is rapidly changing as the target of multimedia systems is moving toward portable applications like laptops, mobiles and iPods etc [15]. These systems highly demand low-power operations, and, thus require low power functional units.

The remaining paper is organized as follows, Section 2 explains about the practical needs, principles and types of image compression, Section 3 explains the Variable Length Encoding Process. Section 4 describes Variable Length Decoding Process, Section 5 includes the Interpretation of Results. Section 6 finally concludes the paper.

II. IMAGE COMPRESSION

A. Practical Needs for Image Compression

The need for image compression becomes apparent when number of bits per image is computed resulting from typical sampling rates and quantization methods [4]. For example, the amount of storage required for given images is (i) a low resolution, TV quality, color video image which has 512×512 pixels/color, 8 bits/pixel, and 3 colors approximately consists of 6×10^6 bits; (ii) A 24×36 mm negative photograph scanned at 12×10^{-6} mm: 3000×2000 pixels/color, 8 bits/pixel, and 3 colors nearly contains 144×10^6 bits; (iii) a 14×17 inch radiograph scanned at 70×10^{-6} mm: 5000×6000 pixels, 12 bits/pixel nearly contains 360×10^6 bits. Thus storage of even a few images could cause a problem. As another example of the need for image compression [15], consider the transmission of low resolution $512 \times 512 \times 8$ bits/pixel $\times 3$ - color video image over telephone lines. Using a 96000 bauds (bits/sec) modem, the transmission would take approximately 11 minutes for just a single image [22], which is unacceptable for most applications.

B. Principles behind Compression

Number of bits required to represent the information in an image can be minimized by removing the redundancy present in it. There are three types of redundancies.

1. Spatial redundancy, which is due to the correlation or dependency between neighboring pixel values.
2. Spectral redundancy, which is due to the correlation between different color planes or spectral bands.
3. Temporal redundancy, which is present because of correlation between different frames in images.

Image compression research [18] aims to reduce the number of bits required to represent an image by removing the spatial and spectral redundancies as much as possible [22].

C. Types of Image Compression

Compression can be divided into two categories [1], Lossless and Lossy compression. In lossless compression schemes, the reconstructed image after compression is numerically identical to the original image. However lossless compression can only achieve a modest amount of compression. Lossless compression is preferred for archival purposes like medical imaging [22], technical drawings, clip art or comics. This is because lossy compression methods, especially when used at low bit rates [21], introduce compression artifacts. An image reconstructed following lossy compression contains degradation relative to the original. Often this is because the compression scheme completely discards redundant information. However the lossy schemes are capable of achieving much higher compression. Lossy methods are especially suitable for natural images such as photos in applications where minor (sometimes imperceptible) loss of fidelity is acceptable to

achieve a substantial reduction in bit rate. The lossy compression is produces imperceptible differences can be called visually lossless.

D. Image Compression using Discrete Cosine Transform

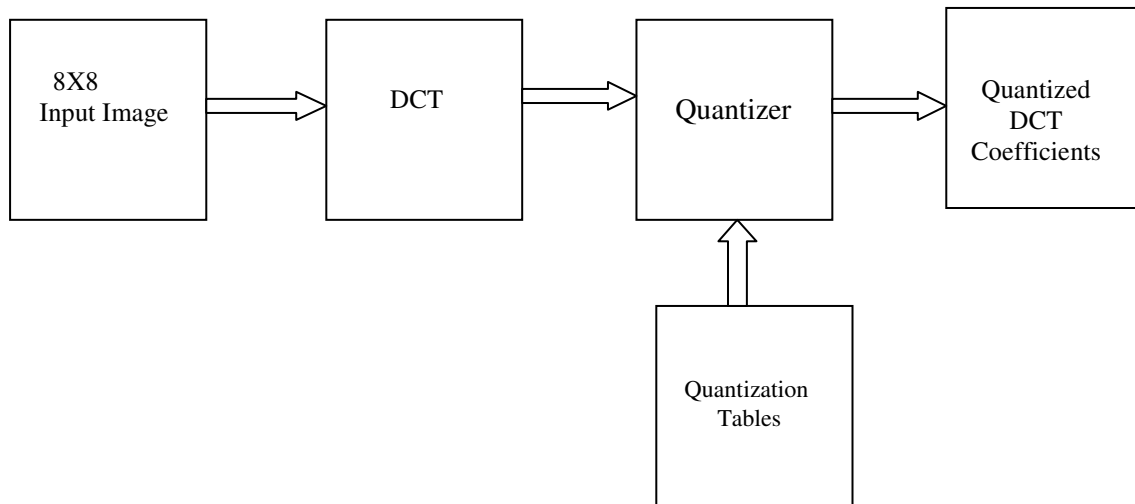


Figure1. Block diagram of DCT Process

In the DCT process input image is divided into non-overlapping blocks of 8×8 pixels [4], and input to the baseline encoder. The pixel values are converted from unsigned integer format to signed integer format, and DCT computation is performed on each block. DCT transforms the pixel data into a block of spatial frequencies that are called the DCT coefficients. Since the pixels in the 8×8 neighbourhood typically have small variations in gray levels, the outputs of the DCT will result in most of the block energy being stored in the lower spatial frequencies [15]. On the other hand, the higher frequencies will have values equal or close to zero and hence, can be ignored during encoding [9] without significantly affecting the image quality. The selection of frequencies based on which frequencies are most important [15] and which ones are less important can affect the quality of the final image.

The selection of quantization values is critical since it affects both the compression efficiency [4], and the reconstructed image quality. High frequency coefficients have small magnitude for typical video data, which usually does not change dramatically between neighbouring pixels. Additionally, the human eye is not as sensitive to high frequencies as to low frequencies [5]. It is difficult for the human eye to discern changes in intensity or colors that occur between successive pixels. The human eye tends to blur these rapid changes into an average hue and intensity. However, gradual changes over the 8 pixels in a block are much more discernible than rapid changes. When the DCT is used for compression purposes, the quantizer unit attempts to force the insignificant high frequency coefficients to zero while retaining the important low frequency coefficients. The 2-D DCT transforms an 8×8 block of spatial data samples into an 8×8 block of spatial frequency components [15]. These DCT coefficients are then used as input to the Variable Length Encoder which will further compress the image data [9]. The compressed image data can be decompressed using Variable Length Decoder [10] and then IDCT transforms spatial frequency components back into the spatial domain [15] to successfully reconstruct the image.

III. VARIABLE LENGTH ENCODING

Variable Length Encoding (VLE) is the final lossless stage of the video compression unit. VLE is done to further compress the quantized image [13]. VLE consists of the following three steps:

- *Zigzag scanning*
- *Run Length Encoding (RLE), and*
- *Huffman coding.*

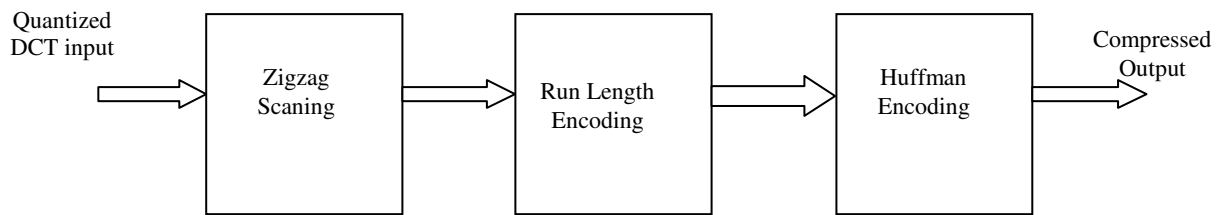


Figure 2. Variable Length Encoder

A. Zigzag Scanning



Figure 3. Block diagram of Zigzag Scanner

The quantized DCT coefficients obtained after applying the Discrete Cosine Transformation to 8×8 block of pixels they are fed as input to the Variable Length Encoder (VLE). These quantized DCT coefficients will have non-zero low frequency components in the top left corner of the 8×8 block and higher frequency components in the remaining places [17]. The higher frequency components approximate to zero after quantization. The low frequency DCT coefficients are more important than higher frequency DCT coefficients. Even if we ignore some of the higher frequency coefficients, we can successfully reconstruct the image from the low frequency coefficients only. The Zigzag Scanner block exploits this property [7]. In zigzag scanning, the quantized DCT coefficients are read out in a zigzag order, as shown in the figure 4. By arranging the coefficients in this manner, RLE and Huffman coding can be done to further compress the data. The scan puts the high-frequency components together. These components are usually zeroes.

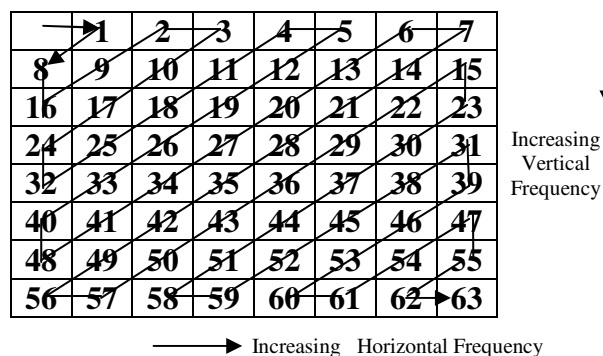


Figure 4. Zigzag Scan Order

Since the zigzag scanning requires that all the 64 DCT coefficients are available before scanning, we need to store the serially incoming DCT coefficients in a temporary memory. For each of the 64 DCT coefficients obtained for each 8×8 block of pixels we have to repeat this procedure. So at a time either scanning is performed or storing of incoming DCT coefficients is done. This will slow down the scanning process. So in order to overcome this problem and to make scanning faster, we propose a new architecture for zigzag scanner. In the proposed architecture, two RAM memories will be used in the zigzag scanner [17]. One of the two RAM memories will be busy in storing the serially incoming DCT coefficients while scanning is performed from the other RAM memory. So except for first 64

clock cycles i.e., until 64 DCT coefficients of first 8x8 blocks become available, the zigzag scanning and storing of serially incoming DCT coefficients is performed simultaneously [11]. So by using two RAM memories we will be able to scan one DCT coefficient in each clock cycle except for first 64 clock cycles.

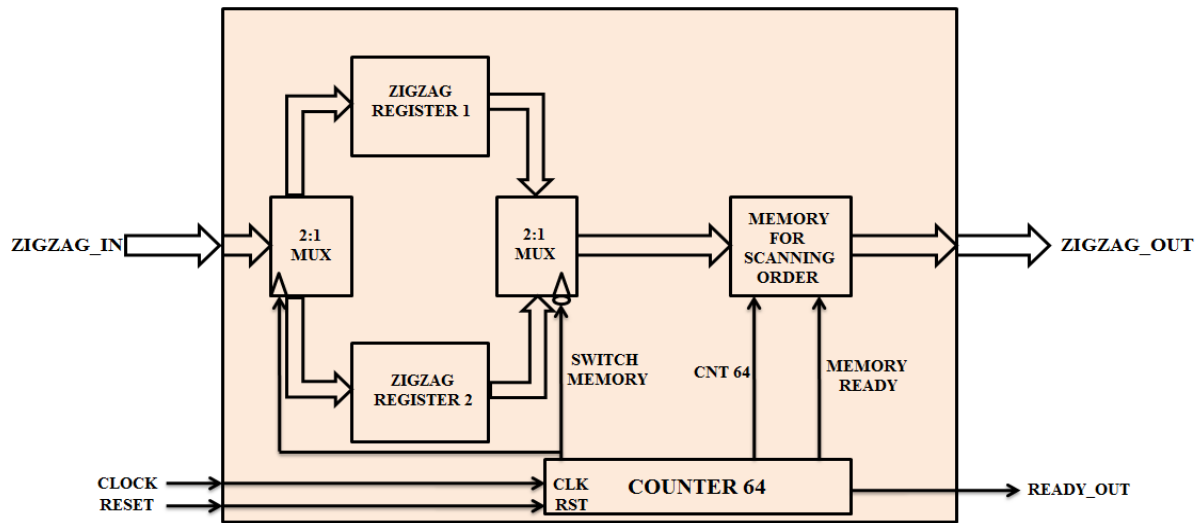


Figure 5. Internal Architecture of the Zigzag Scanner

B. Run Length Encoding (RLE)

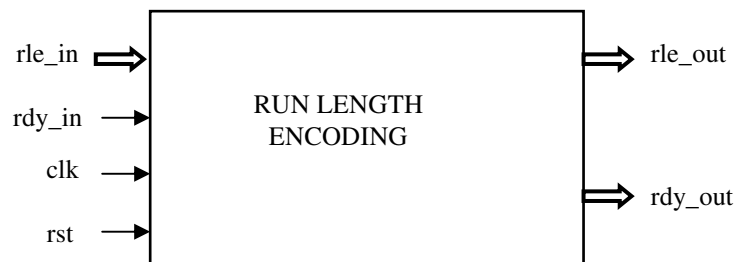


Figure 6. Block diagram of Run- Length Encoder

The quantized coefficients are read out in a zig-zag order from DC component to the highest frequency component. RLE is used to code the string of data from the zig-zag scanner. The conventional Run length encoder codes the coefficients in the quantized block into a run length (or number of occurrences) and a level or amplitude. For example, transmits four coefficients of value "10" as: {10,10,10,10}. By using RLE [8], the level is 10 and the run of a value of 10 is four. By using RLE, {4,10} is transmitted, reducing the amount of data transmitted. Typically, RLE [10] encodes a run of symbols into two bytes, a count and a symbol. By defining an 8 x 8 block without RLE, 64 coefficients are used. To further compress the data, many of the quantized coefficients in the 8 x 8 block are zero. Coding can be terminated when there are no more non-zero coefficients in the zig-zag sequence [9]. Using the "end-of-block" code terminates the coding.

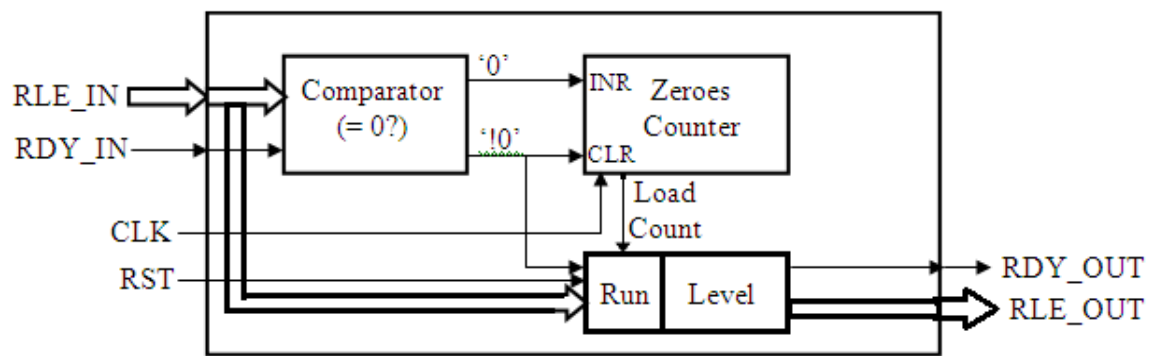


Figure 7. Internal architecture of run-length encoder.

But normally in a typical quantized DCT matrix the number of zeroes is more [5] compared to non-zero coefficients being repeated [4]. So in the proposed architecture for run-length encoder we exploit this property of more number of zeroes in the DCT matrix. In the proposed architecture the number of intermediate zeros in between non-zero DCT coefficients are counted [10] unlike the conventional run-length encoder where number of occurrences of repeated symbols are counted.

Conventional RLE	Proposed RLE
31	31
1,0	1,1
1,1	1,2
1,0	0,1
1,2	5,2
1,1	EOB
5,0	
1,1	
EOB	

Figure 8. Comparison between conventional and proposed RLE

C. Huffman encoding

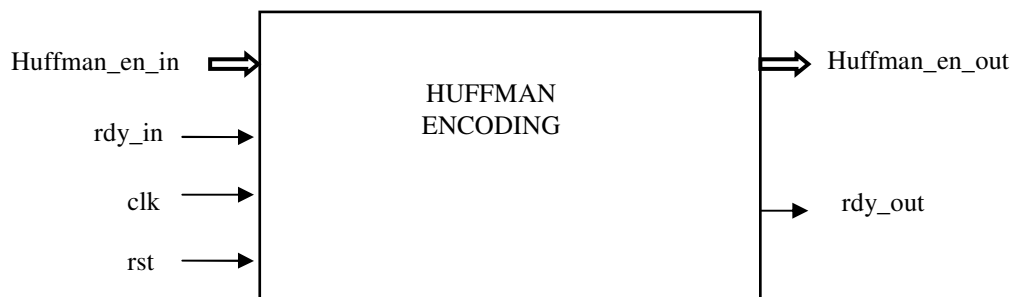


Figure 9. Block diagram of Huffman Encoder

Huffman coding is used to code values statistically according to their probability of occurrence [11]. Short code words are assigned to highly probable values and long code words to less probable values. The procedure for Huffman coding involves the pairing of run/value combinations. The input run/value combinations are written out in the order of decreasing probability. The run/value combination with the highest probability is written at the top, the least probability is written down last. The least two probabilities are then paired and added. A new probability list is then formed with one entry as the previously added pair. The least run/value combinations in the new list are then paired. This process is continued till the list consists of only one probability value. The values "0" and "1" are arbitrarily assigned to each element in each of the lists.

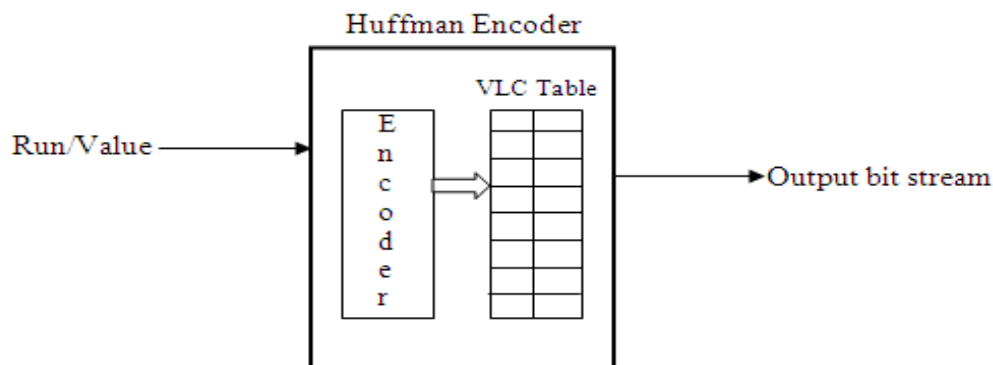


Figure 10. Internal architecture of Huffman encoder

In the proposed architecture, the Huffman encoding is done by making use of a lookup table [3]. The lookup table is formed by arranging the different run-value combinations in the order of their probabilities of occurrence with the corresponding variable length Huffman codes [6]. When the output of the run-length encoder in form of run-value combination is fed to the Huffman encoder, the run-value combination received will be searched in the lookup table, when run-value combination is found its corresponding variable length Huffman code [14] will be sent to output. This approach of designing Huffman encoder not only simplifies the design but also results in less power consumption [6]. Since we are using lookup table approach, the only part of encoder corresponding to the current run-length combination will be active and other parts of the encoder will not be using any power. So turning off the inactive components of a circuit in the Huffman encoder, results in less power consumption.

IV. VARIABLE LENGTH DECODING

The variable length decoder is the first block on the decoder side. It decodes the variable length encoder output to yield the quantized DCT coefficients [10]. The variable length decoder consists of three major blocks, namely,

1. Huffman Decoding.
2. Run-length Decoding.
3. Zigzag Inverse Scanning.

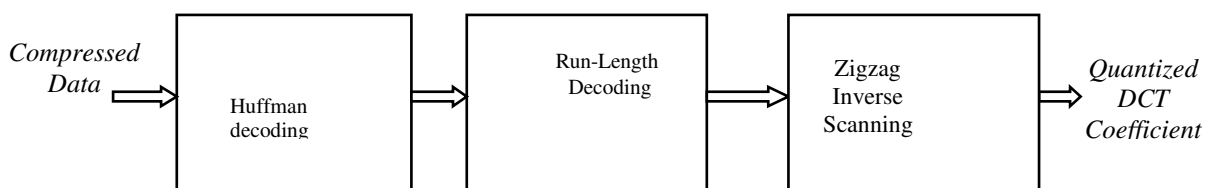


Figure11. Variable Length Decoder

A. Huffman Decoding

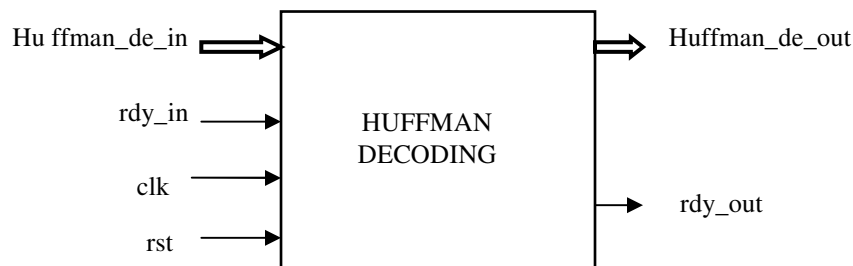


Figure 12. Block diagram of Huffman decoder

The huffman decoder forms the front end part of the variable length decoder. The internal architecture of the huffman decoder is same as the huffman encoder. The same VLC huffman coding table which was used in the huffman encoder is also used in the huffman decoder [17]. The input encoded data is taken and a search is done for the corresponding run/value combination in the VLC table. Once the corresponding run/value combination [14] is found, it is sent as output and huffman starts decoding next coming input.

The VLC huffman coding table which we are using in both the huffman encoder and the huffman decoder, reduces the complexity of the huffman decoder. It is not only reduces the complexity [7] but also reduces the dynamic power in the huffman decoder, since only the part of the circuit is active at a time.

B. FIFO

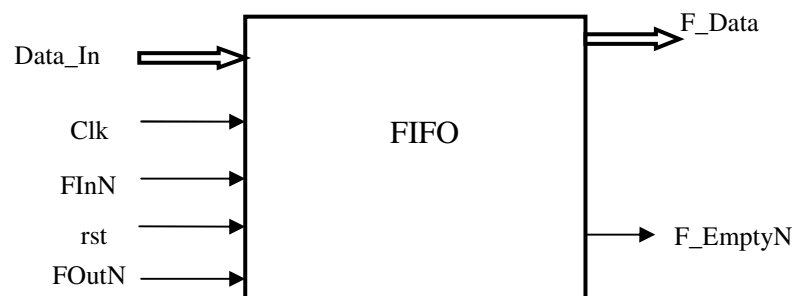


Figure 13. Block diagram of FIFO

The First In First Out (FIFO) also forms the part of the decoder part, The FIFO is used between the huffman decoder and the run-length decoder. The FIFO is used to match the operating speed between the huffman decoder and run-length decoder [11]. The huffman decoder sends a decoded output to the run-length decoder in the form of run/value combination. The run-length combination takes this as input and starts decoding [12]. Since here the run in the run/value combination represents the number of zeroes in between consecutive non-zero coefficients, the zero '0' is sent as output for next 'run' number of clock cycles. Until then the run-length decoder can't accept other run/value combination. And we know that the huffman decoder decodes one input to one run/value combination in every clock cycle. So huffman decoder can't be connected directly to run-length decoder. Otherwise the run-length decoder can't decode correctly. So to match the speed between the huffman decoder and run-length decoder, the FIFO is used. The output of the huffman decoder is stored onto the FIFO, the run-length decoder takes one decoded output of huffman decoder from the FIFO when it is finished with decoding of the present input to it. So after run-length decoder finishes decoding of the present input, it has to send a signal to the FIFO to feed it a new input. This signal is sent to the FOutN pin, which is read out pin of the FIFO. The FInN pin is used to write onto FIFO, the huffman decoder generates the signal for this while it has to write a new input onto the FIFO. So the FIFO acts as a synchronizing device between the huffman decoder [9] and the run-length decoder.

C. Run-Length Decoder

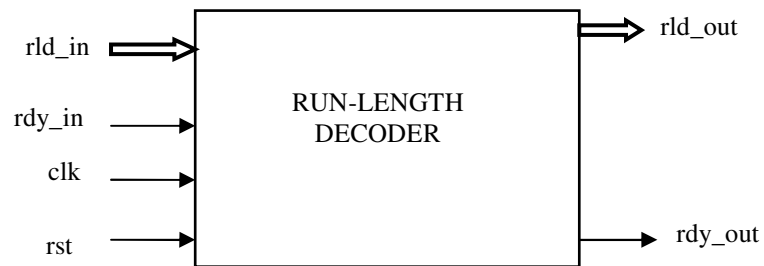


Figure14. Block diagram of Run- Length Decoder

The Run-length decoder forms the middle part of the variable length decoder [10]. It takes decoded output from huffman decoder through FIFO. When the huffman decoder decodes one input and stores the decoded output onto the FIFO, then the FIFO becomes non-empty (the condition when at least one element is stored on the FIFO). The FIFO then generates the signal F_EmptyN. This signal is used as rdy_in signal to the run-length decoder. So when huffman decoder decodes one input and stores it onto the FIFO[7], then a ready signal is generated to the run-length decoder to initiate the decoding process.

The run-length decoder takes the input in the form of a run/value combination, then separates run and value parts. The run here represents number of zeroes to output before sending out the non-zero level 'value' in the run/value combination. So for example if {5,2} is input to the run-length decoder then it sends 5 zeroes (i.e., 5 '0') before transmitting the non-zero level '2' to the output. Once the run-length decoder sends out a non-zero level, then it means that it is finished with the decoding of the present run/value combination, and it is ready for the next run/value combination. So for this it generates the rdy_out signal to the FIFO, to indicate that it has finished decoding of present input and ready for decoding the next run/value combination. This rdy_out is connected to the FOutN pin of the FIFO, which is read out pin of the FIFO[16]. Upon receiving this signal the FIFO sends out a new run/value combination to the run-length decoder, to initiate run-length decoding process for the new run/value combination.

D. Zigzag Inverse Scanner

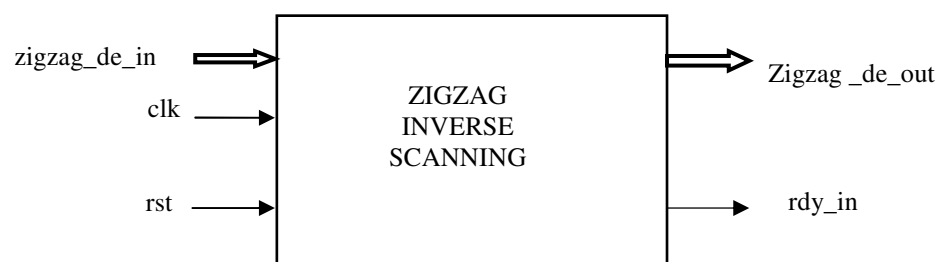


Figure15. Block diagram of Zigzag Scanner

The zigzag inverse scanner forms the last stage in the variable length decoder [17]. The working and architecture, everything is similar to the zigzag scanner, except that the scanning order will be different. The zigzag inverse scanner gets the input from the run-length decoder, starts storing them in one of the two RAMs [13], until it receives all 64 coefficients. Once it receives all the 64 coefficients, it starts inverse scanning to decode back the original DCT coefficients. Meanwhile the incoming DCT coefficients are getting stored in another RAM [2]. Once scanning from one RAM is finished, it starts scanning from another RAM and meanwhile the incoming DCT coefficients gets stored in first RAM [5]. So this process is repeated until all the DCT coefficients are scanned. There will be delay of 64 clock cycles before the output appears. Once after that, for every clock cycle an element will scanned continuously.

V. INTERPRETATION OF RESULTS

1.Simulation Results

A. Zigzag Scanning

The simulation of zigzag scanning is done using the following test sequence.

31	0	1	0	0	0	0	0	1	2	0	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11		62	63

The figure 16 shows the simulation waveform of the zigzag encoding block for the above test sequence.

B. Run Length Encoding

The simulation of the run-length encoding block is done using the output sequence obtained in zigzag scanning process, which appears at the input of the run-length encoder as below.

31	0	1	0	2	1	0	0	0	0	0	2	0	0
0	1	8	16	9	2	3	10	17	24	32	25		62	63

The figure 17 shows the simulation waveform of the run-length encoding block for the above test input sequence.

C. Huffman Encoding

31	1,1	1,2	0,1	5,2	EOB
----	-----	-----	-----	-----	-----

The output of the run-length encoder is used as the test sequence to the Huffman encoder. The output of the run-length encoder will appear as below. The figure 18 shows the simulation waveform of the Huffman encoding block for the above test input sequence.

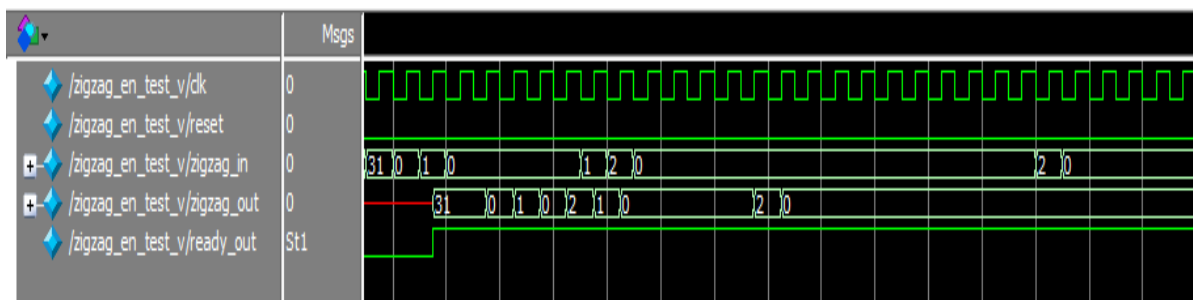


Figure16. Simulation waveform of Zigzag Scanning

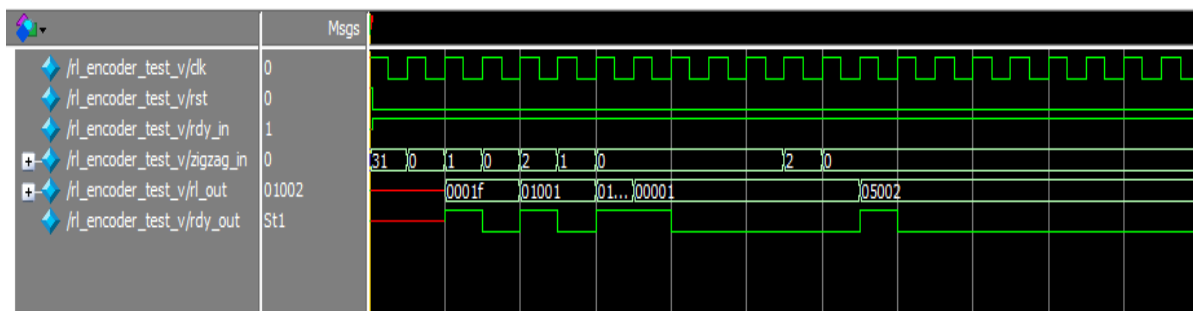


Figure17. Simulation waveform of Run-length Encoder

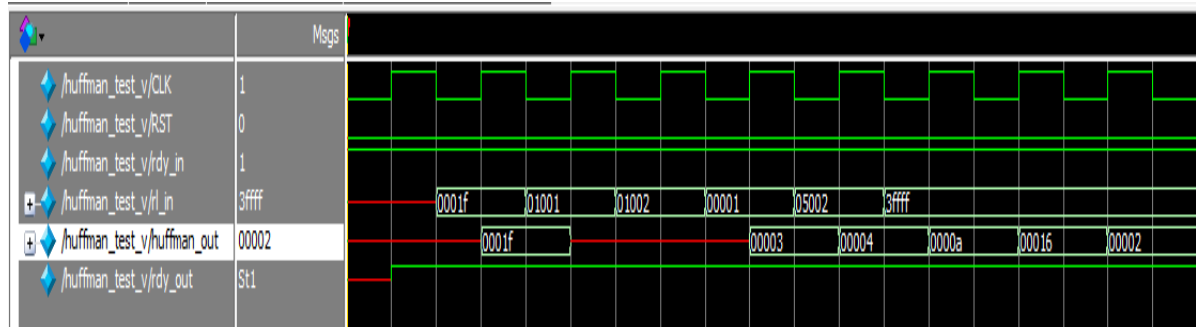


Figure18. Simulation waveform of Huffman Encoder

D. Huffman Decoding

The compressed data from the variable length encoder is fed to the Huffman decoding block. The output obtained from the Huffman encoding block is used as a test bench to the Huffman decoding block. The fig 19 shows the simulation waveform for the Huffman decoding block.

E. Run Length Decoding

After Huffman decoder decodes the compressed bitstream, the decoded input is fed to run-length decoder block. The fig 20 shows the simulation waveform of run-length decoder.

F. Zigzag Inverse Scanner

The output of the run-length decoder is given as input to the zigzag inverse scanner, which will output the quantized DCT coefficients. The fig 21 shows the simulation waveform of Zigzag Inverse Scanner.

2. Synthesis Results

Once all the behavioral verifications (simulation) are done, syntheses of blocks are performed. The Cadence RTL compiler is used for synthesis and design is implemented on 90nm standard cells. The layout is generated using VLSI Backend EDA tool. After synthesis of each blocks the area and power analysis are performed for each of the block used in the design. The Layout of the Run Length Encoder is shown in the figure 22.

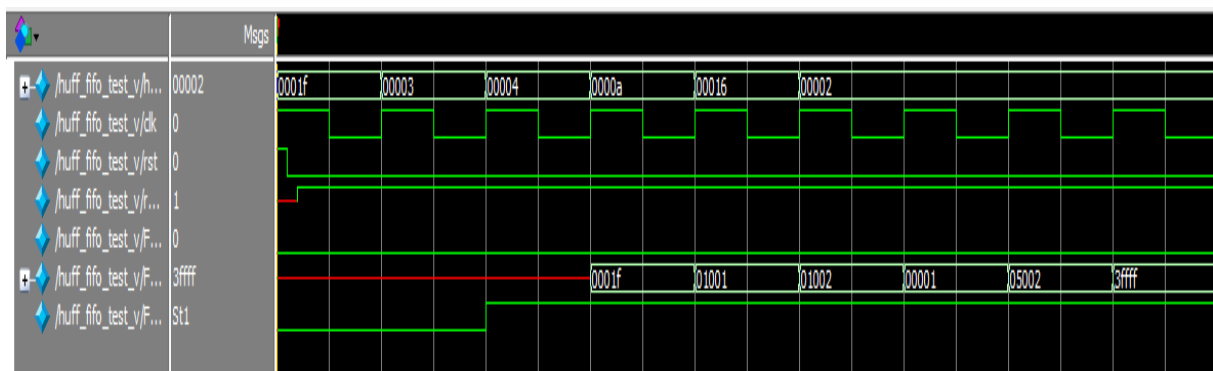


Figure19. Simulation waveform of Huffman Decoder

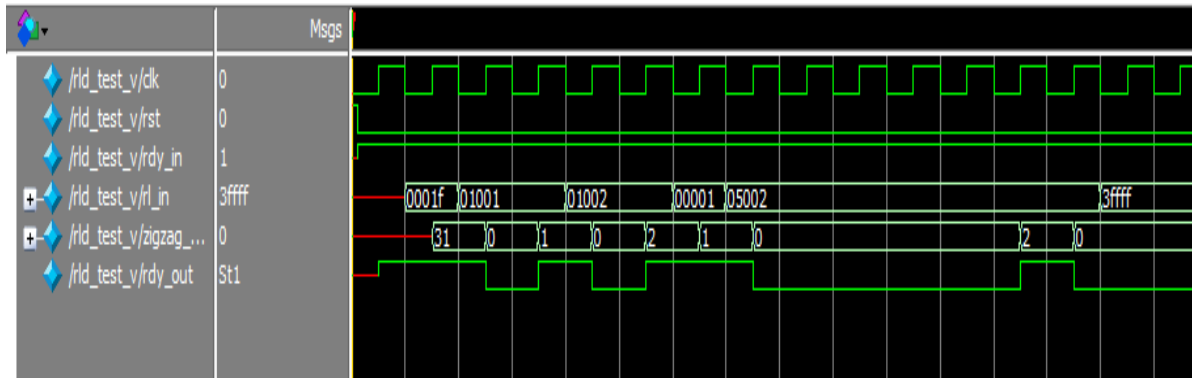


Figure20. Simulation waveform of Run-length Decoder

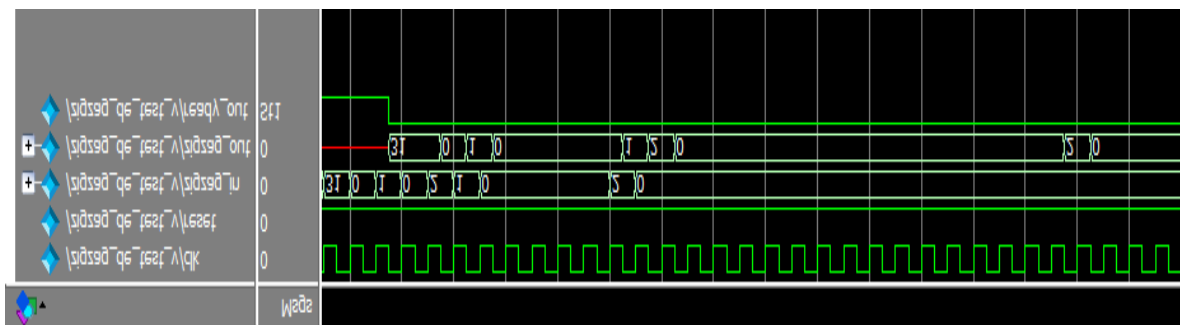


Figure21. Simulation waveform of Zig-Zag Inverse Scanner

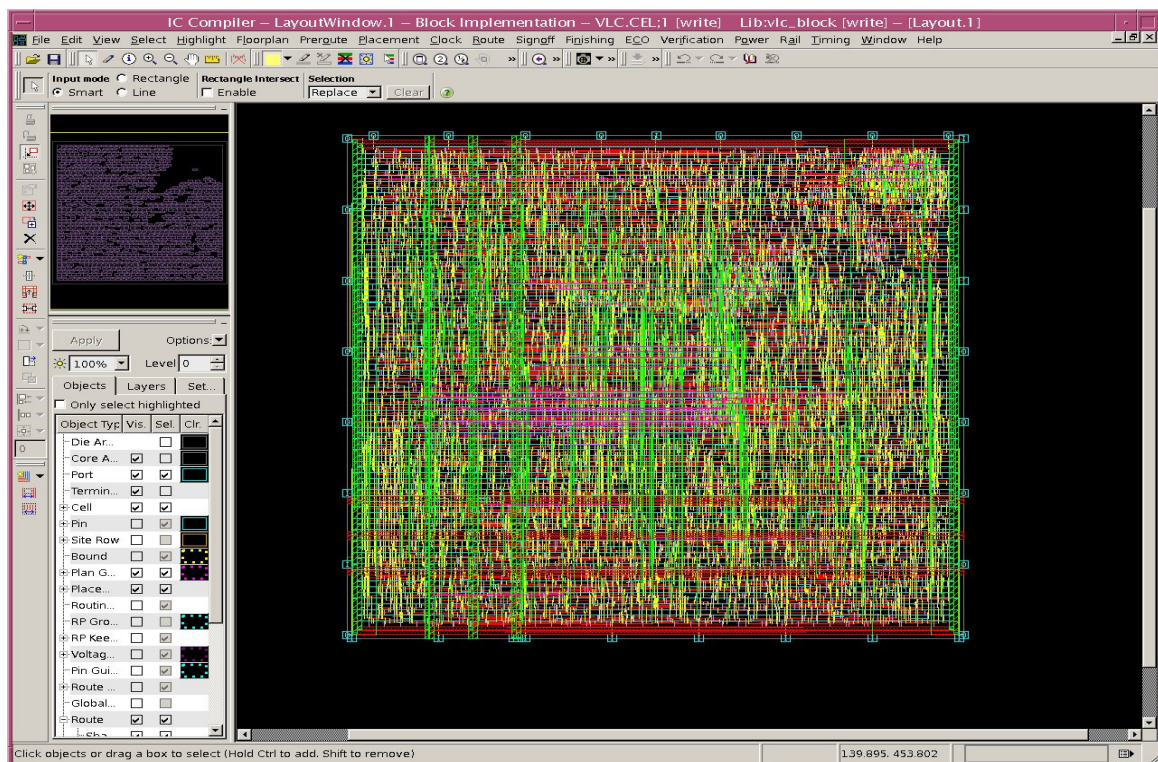
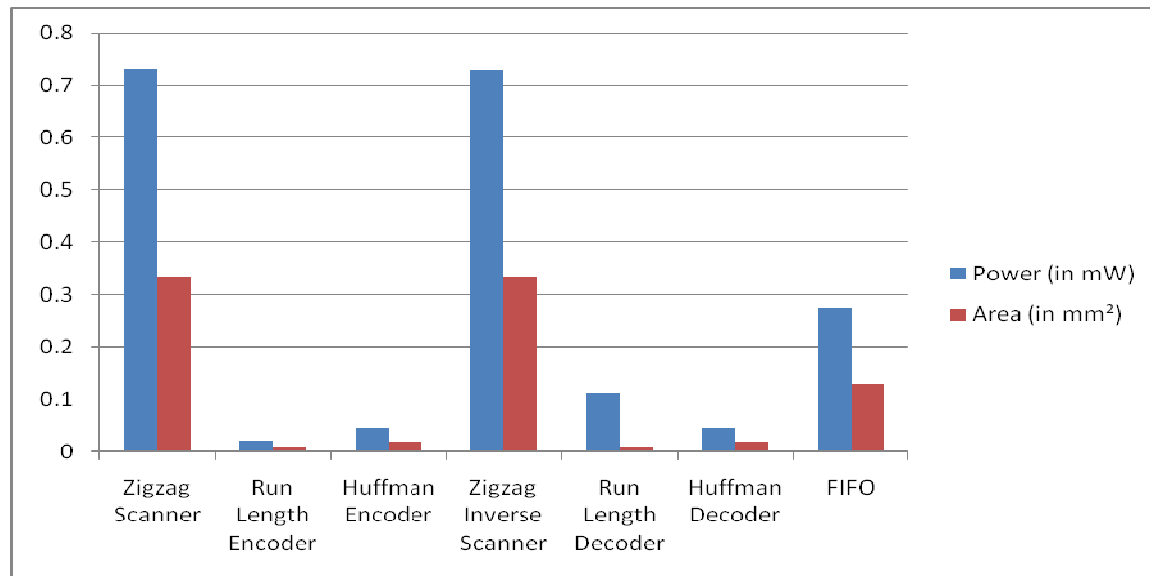


Figure22. Layout of Variable Length Encoder

The power and area characteristics of each of the blocks in the design are tabulated as shown in Table 1, and the characteristics are represented in the form of graph are also show in graph 1.

Table 1: Power & Area Characteristics

Designs \ Features	Power(in mW)	Area
Zigzag Scanner	0.7319	0.3349
Run Length Encoder	0.0208	0.0103
Huffman Encoder	0.0451	0.0171
Zigzag Inverse Scanner	0.7285	0.3359
un Length Decoder	0.1110	0.0071
Huffman Decoder	0.0451	0.0171
FIFO	0.2744	0.1287

**Figure 23.** Representation of Power & Area Characteristics

3. Power Comparisons

The power comparison of the proposed Architecture is as shown below.

3.1 Power comparison of Huffman Decoder

Table 2: Power Comparison for Huffman decoders

Table Size	Huffman decoder type	Power (in μ W)
100	Power Analysis of the Huffman Decoding Tree, by Jason McNeely [6]	95
	Proposed Architecture	45

3.2 Power comparison of RL- Huffman Encoder Combination

Table 3: Power Comparison for RL-Huffman Encoders

RL-Huffman Encoding Type	Power (in μ W)
RL-Huffman Encoding for Test Compression and Power Reduction in Scan Applications [8]	90
Proposed Architecture	65.9

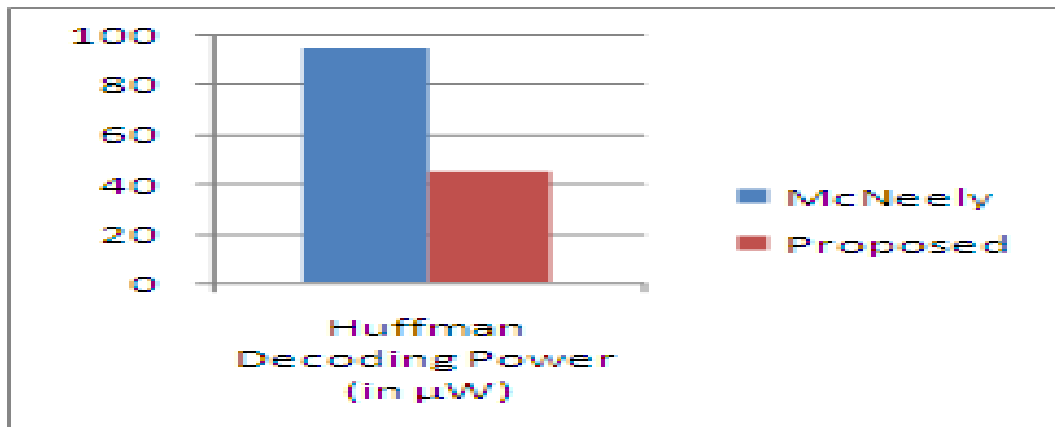


Figure 24. Power Comparison for Huffman decoders

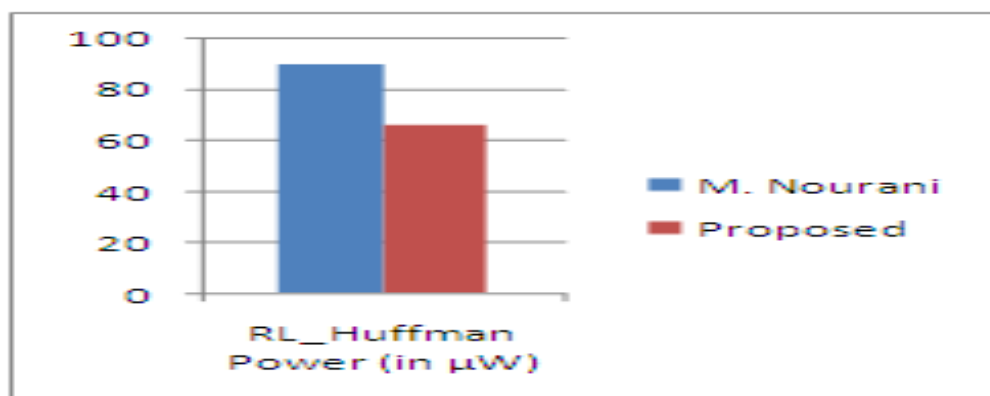


Figure 25. Power Comparison for RL-Huffman

3.3 Percentage of Power Saving

The percentage of power savings from proposed design are calculated and are tabulated as shown in following table 4.

Table 4: Percentage of Power Savings

Encoding Type	Comparison with	Percentage of Power Saving
Huffman decoder	Power Analysis of the Huffman Decoding Tree [6]	52.63%
RL-Huffman Encoder	RL-Huffman Encoding for Test Compression and Power Reduction in Scan Applications [8]	26.77%

VI. CONCLUSION

In this paper we described a Low Power Architecture for Variable Length Encoding and Variable Length Decoding for image processing applications. The designing and modeling of the all the blocks in the design is done by using synthesizable Verilog HDL with Low Power approach. The proposed architecture is synthesized using RTL compiler and it is mapped using 90 nm standard cells. The Simulation of all the blocks in the design was done using Modelsim. A detailed analysis for power and area was done using RTL compiler from CADENCE. Power consumptions of variable length encoder and decoder are limited to 0.798mW and 0.884mW with minimum area. A 53% power saving is achieved in the dynamic power of huffman decoding [6] by including the lookup table approach and also a 27% of power saving is achieved in the RL-Huffman encoder [8].

REFERENCES

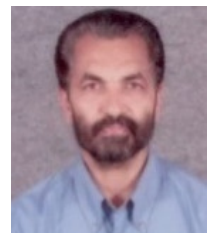
- [1] G. Wallace, "The JPEG still-image compression standard," Commun. ACM, vol. 34, pp. 30–44, Apr. 1991. ACM, vol. 34, pp. 30–44, Apr. 1991.
- [2] *Low Power Look-Up Tables for Huffman Decoding*, by Jason McNeely and Magdi Bayoumi, The Centre of Advanced Computer Studies, University of Louisiana, Lafayette.
- [3] *Direct Huffman Coding and Decoding using the Table of Code-Lengths*, by Reza Hashemian, Senior Member, IEEE Northern Illinois University DeKalb, Illinois 60115, USA.
- [4] *A Novel Fast DCT Coefficient Scan Architecture*, by Da An, Xin Tong, Bingqiang Zhu and Yun He, State Key Laboratory on Microwave and Digital Communications Tsinghua National Laboratory for Information Science and Technology Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.
- [5] A Novel VLSI Architecture for Image Compression using DCT and Quantization. By Vijaya Prakash A M and K. S Gurumurthy, IJCSNS Vol .10 No.9 September 2010.
- [6] "Power Analysis of the Huffman Decoding Tree", Jason McNeely, Yasser Ismail, Magdy A. Bayoumi and Peiyi Zaho. The Center for Advanced Computer Studies, University of Louisiana at Lafayette. IEEE Conference on Image Processing (ICIP) California USA October 2008.
- [7] D.A. Huffman, "A method for construction of minimum-redundancy codes", Proc. IRE, Vol. 40, pp. 1098-1101, Sept. 1952.
- [8] "RL-Huffman Encoding for Test Compression and Power Reduction in Scan Applications". Mehrdad Nourani and Mohammed H. Tehranipour Center for Integrated Circuits and Systems, The University of Texas at Dallas. ACM Transactions on Design Automation of Electronic Systems vol.10.No1. Jan 2005.
- [9] *Joint Optimization of Run-Length Coding, Huffman Coding, and Quantization Table with Complete Baseline JPEG Decoder Compatibility*, by En-hui Yang, Fellow, IEEE, and Longji Wang, Member, IEEE.
- [10] *A Low-Power Variable Length Decoder for MPEG-2 Based on Successive Decoding of Short Codewords* Sung-Won Lee, Student Member, IEEE, and In-Cheol Park, Senior Member, IEEE.
- [11] *A Fast Parallel Huffman Decoder for FPGA Implementation* ACTA TECHNICA NAPOCENSIS Electronics and Telecommunications Volume 49, Number 1, 2008.
- [12] 'JPEG Architecture and Implementation Issues', by Dr. Sri Krishnan Department of Electrical and Computer Engineering Ryerson University.
- [13] A Study and Implementation of the Huffman Algorithm Based on Condensed Huffman Table, by Bao Ergude, Li Weisheng, Fan Dongrui, Ma Xiaoyu, School of Software, Beijing Jiaotong University; and Key Laboratory of Computer System and Architecture, Institute of Computer Technology, Chinese Academy of Sciences, 2008.
- [14] An efficient Memory Construction Scheme for an Arbitrary Side Growing Huffman table, by Sung-Wen Wang, Shang-Chih Chuang, Chih-Chieh Hsiao, Yi-Shin Tung and Ja-ling Wu CMLab, Dept. CSIE NTU, Setabox Corporation, Graduate Institute of Networking and Multimedia NTU, NOVEMBER 2008.
- [15] A Novel VLSI Architecture for Image Compression Model using Low Power DCT. By Vijaya Prakash A M and K. S Gurumurthy, WASET Vol .72 December 2010 Singapur.
- [16] Balance of 0, 1 Bits for Huffman and Reversible Variable-Length Coding, by Jia-Yu Lin, Ying Liu, and Ke-Chu Yi, MARCH 2004.
- [17] *Parallel Zigzag Scanning and Huffman Coding for a GPU-Based MPEG-2 Encoder*, by Pablo Montero, Javier Taibo Videalab University of A Coruna, A Coruna, Spain. Victor Gulias, MADS Group University of A Coruna, A Coruna, Spain. Samuel Rivas LambdaStream S.L. A Coruna, Spain.
- [18] Discrete Wavelet Transform for Image Compression and A Model of Parallel Image Compression Scheme for Formal Verification...Proceedings of the World Congress on Engineering 2007 Vol I by Kamrul Hasan Talukder and Koichi Harada.
- [19] Adaptive Context Based Coding for Lossless Color Image Compression, by Yongli Zhu and Zhengya XU School of Computer Science and Technology North China Electrical Power University. IMACS Conference on 'Computational Engineering in Systems Applications' (CESA) 2006 Beijing China.
- [20] An Efficient and Selective Image Compression Scheme using Huffman and adaptive Interpolation by Sunil Bhushan and Shipra Sharma. 24th International Conference Image and Vision Computing New Zealand (IVCNZ 2009).
- [21] Low Bit Rate Image Coding Based on Wavelet Transform and Color Correlative Coding by Wenna Li and Zhaohua Cui "2010 International Conference on Computer design and Applications (ICCD A 2010).
- [22] Medical Image Coding based on Wavelet transform and Distributed arithmetic Coding by Li Wenna ,

Authors

Vijaya Prakash A.M Obtained is B.E from UVCE Bangalore of Bangalore University in the year 1992, Post graduated in M.E from SDMCET Dhawad of Karnataka University in the year 1997 and presently pursuing Ph.D. from Dr. M. G. R University Chennai. He has been actively guiding PG and UG student's projects in the area of VLSI Design and Image Processing. He has around 8 technical paper publications in international journals and international conferences. Currently he has been working as Associate Professor in Electronics and Communication Engineering Department, Bangalore Institute of Technology Bangalore-04. He has presented paper in Singapur. His research interests are Low Power VLSI, Image Processing, Retiming and Verification of VLSI Design, Synthesis and Optimization of Digital Circuits. He is a member of IMAPS and ISTE.



K.S Gurumurthy obtained his B.E degree from M.C.E – Hassan of Mysore University in the year 1973. He got his M.E Degree from University of Roorkee (now IIT-Roorkee) in 1982. He joined UVCE in 1982 and he has since been teaching Electronics related subjects. He obtained his Ph.D. degree in 1990 from IISc Bangalore. Presently he is a professor in the DOS in E & CE, UVCE, BU, Bangalore-1. He is a “University gold medal” winner from University of Roorkee and a recipient of the “Khosla award” for the best technical paper published in 1982. He has successfully guided 4 Ph.D., 2 M.Sc-Engg (by research) candidates and guided a number of UG and PG projects. He has around 75 technical paper publications in journals and international conferences. He has presented papers in JAPAN, FRANCE, MALAYASIA and USA. His interests are Low power VLSI, Multi valued logic circuits, Deep submicron Devices. He is a member of IEEE and ISTE.



VERIFICATION ANALYSIS OF AHB-LITE PROTOCOL WITH COVERAGE

Richa Sinha¹, Akhilesh Kumar² and Archana Kumari Sinha³
^{1&2}Department of E&C Engineering, NIT Jamshedpur, Jharkhand, India
³Department of Physics, S.G.G.S. College, Patna City, Bihar, India

ABSTRACT

The SoC design faces a gap between the production capabilities and time to market pressures. The design space, grows with the improvements in the production capabilities in terms of amount of time to design a system that utilizes those capabilities. On the other hand shorter product life cycles are forcing an aggressive reduction of the time-to-market. Fast simulation capabilities are required for coping with the immense design space that is to be explored; these are especially needed during early stages of the design. This need has pushed the development of transaction level models, which are abstract models that execute dramatically faster than synthesizable models. The pressure for fast executing models extends especially to the frequently used and reused communication libraries. The presents paper describes the system level modelling of the Advanced High-performance Bus Lite (AHB-Lite) subset of AHB which part of the Advanced Microprocessor Bus Architecture (AMBA). The work on AHB-Lite slave model, at different test cases, describing their simulation speed. Accuracy is built on the rich semantic support of a standard language SystemVerilog on the relevant simulator Riviera has been highlighted.

KEYWORDS: AMBA(Advanced Microcontroller Bus Architecture), AHB-Lite(Advanced High performance Bus-Lite), SystemVerilog, SoC(System on chip), Verification intellectual property (VIP).

I. INTRODUCTION

The bus protocol used by the CPU is an important aspect of co-verification since this is the main communication between the CPU, memory, and other custom hardware. The design of embedded systems in general and a SoC in special will be done under functional and environmental constraints. Since the designed system will run under a well-specified operating environment, the strict functional requirements can be concretely defined. The environment restrictions on the other hand are more diverse: e.g. minimizing the cost, footprint, or power consumption. Due to the flexibility of a SoC design, ARM processors use different bus protocols depending on when the core was designed for achieving the set goals, involves analyzing a multi-dimensional design space. The degrees of freedom stem from the process element types and characteristics, their allocation, the mapping of functional elements to the process elements, their interconnection with busses and their scheduling. The enormous complexity of these protocol results from tackling high-performance requirements. Protocol control can be distributed, and there may be non-atomicity or speculation.

AHB-Lite systems based around the Cortex-M™ processors ARM delivers the DMA-230 "micro" DMA controller [13]. ARM delivers DMA controllers for both high-end, high-performance AXI systems based on the Cortex-A™ and Cortex-R™ families and cost-efficient AHB systems built around Cortex-M™ and ARM9 processors.

The CoreLink Interconnect family includes the following products for AMBA protocols:

- Network Interconnect (NIC-301) for AMBA 3 systems including support for AXI, AHB and APB
- Advanced Quality of Service (QoS-301) option for NIC-301

The third generation of AMBA defines the targeted at high performance, high clock frequency system designs and includes features which make it very suitable for high speed sub-micrometer interconnect. In the present paper the some discussion is made on the family of AMBA and a small introduction on SystemVerilog language which used during VIP. And also the briefly described the AHB-Lite Protocol. Further verification intellectual property (VIP) of slave of the AHB-Lite protocol with different test cases is shown.

II. AMBA PROTOCOLS

Figure 1. shows the different protocols performances from the time of initialization[9].

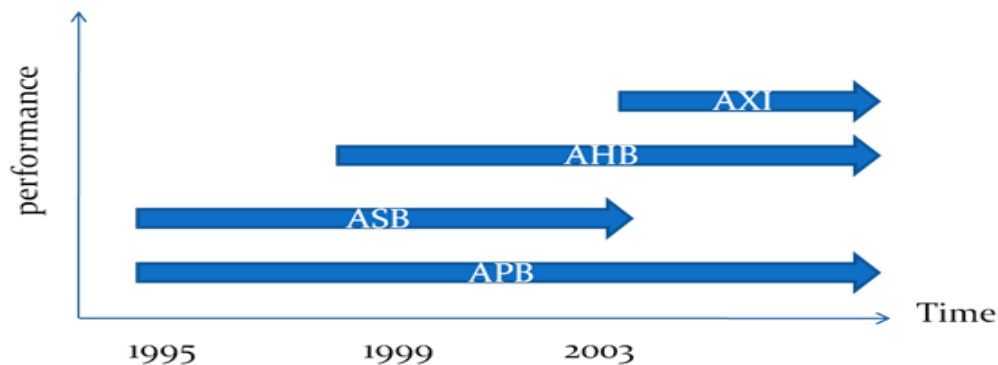


Figure 1. Protocols of AMBA[9]

- ✚ APB (Advanced Peripheral Bus) mainly used as an ancillary or general purpose register based peripherals such as timers, interrupt controllers, UARTs, I/O ports, etc. It is connected to the system bus via a bridge, helps reduce system power consumption. It is also easy to interface to, with little logic involved and few corner- cases to validate.
- ✚ AHB (Advanced High Performance Bus) is for high performance, high clock frequency system modules with suitable for medium complexity and performance connectivity solutions. It supports multiple masters.
- ✚ AHB-Lite is the subset of the full AHB specification which intended for use where only a single bus master is used and provides high-bandwidth operation.

III. SYSTEMVERILOG

SystemVerilog is a Hardware Description and Verification Language based on Verilog. Although it has some features to assist with design, the thrust of the language is in verification of electronic designs. The bulk of the verification functionality is based on the Open Vera language donated by Synopsys[12]. SystemVerilog has just become IEEE standard P1800-2005. SystemVerilog is an extension of Verilog-2001; all features of that language are available in SystemVerilog i.e Verilog HDL, VHDL, C, C++.

IV. AHB-LITE PROTOCOL SYSTEM

AMBA AHB-Lite protocol addresses the requirements of high-performance synthesizable designs. It is a bus interface that supports a single bus master and provides high-bandwidth operation.

AHB-Lite implements the features required for high-performance, high clock frequency systems including: [1]

- burst transfers
- single-clock edge operation
- non-tristate implementation

- Wide data bus configurations, 64, 128, 256, 512, and 1024 bits.

The most common AHB-Lite slaves are internal memory devices, external memory interfaces, and high bandwidth peripherals. Although low-bandwidth peripherals can be included as AHB-Lite slaves, for system performance reasons they typically reside on the AMBA Advanced Peripheral Bus (APB). Bridging between this higher level of bus and APB is done using a AHB-Lite slave, known as an APB bridge.

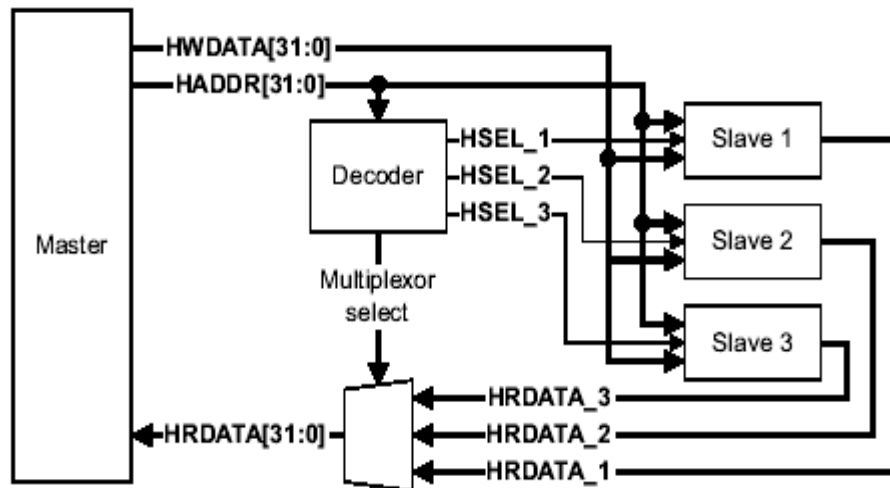


Figure 2. AHB-Lite block diagram

Figure 2. shows a single master AHB-Lite system design with one AHB-Lite master and three AHB-Lite slaves. The bus interconnect logic consists of one address decoder and a slave-to-master multiplexor. The decoder monitors the address from the master so that the appropriate slave is selected and the multiplexor routes the corresponding slave output data back to the master. The main component types of an AHB-Lite system are described in:

- Master
- Slave
- Decoder
- Multiplexor

4.1 Operations of AHB-Lite

The master starts a transfer by driving the address and control signals. These signals provide information about the address, direction, width of the transfer, and indicate if the transfer forms part of a burst. Transfers can be:[11]

Table 1. Transfer type values

Cycle Type	Description	HTRANS[1:0]
IDLE	No bus activity	00
BUSY	Master inserting wait states	01
NON-SEQUENTIAL	Transfer with address not related to the previous transfer	10
SEQUENTIAL	Transfer with address related to the previous transfer	11

The write data bus moves data from the master to a slave, and the read data bus moves data from a slave to the master. Every transfer consists of:[2]

- **Address phase** one address and control cycle
- **Data phase** one or more cycles for the data.

A slave cannot request that the address phase is extended and therefore all slaves must be capable of sampling the address during this time. However, a slave can request that the master extends the data phase by using **HREADY**. This signal when LOW, causes wait states to be inserted into the transfer and enables the slave to have extra time to provide or sample data.

The slave uses **HRESP** to indicate the success or failure of a transfer.

Table 2.Response type values

Description	HRESP[1:0]
Completed Successfully	00
Error occurred	01
Master should retry	10
Perform Split Protocol	11

V. SPECIFICATION DIFFERENT FROM AHB

The AHB-Lite specification differs from the full AHB specification in the following ways[2]:

- Only one master. There is only one source of address, control, and write data, so no Master-to-Slave multiplexor is required.
- No arbiter. None of the signals associated with the arbiter are used.
- Master has no **HBUSREQ** output. If such an output exists on a master, it is left unconnected.
- Master has no **HGRANT** input. If such an input exists on a master, it is tied HIGH.
- Slaves must not produce either a Split or Retry response.
- The AHB-Lite lock signal is the same as **HMASTLOCK** and it has the same timing as the address bus and other control signals. If a master has an **HLOCK** output, it can be retimed to generate **HMASTLOCK**.
- The AHB-Lite lock signal must remain stable throughout a burst of transfers, in the same way that other control signals must remain constant throughout a burst.

VI. COMPATIBILITY

Table 3 shows how masters and slaves designed for use in either full AHB or AHB-Lite can be used interchangeably in different systems.

Table 3

Component	Full AHB system	AHB-Lite system
Full AHB master	✓	✓
AHB-Lite master	Use standard AHB master wrapper	✓
AHB slave (no Split/Retry)	✓	✓
AHB slave with Split/Retry	✓	Use standard AHB master wrapper

VII. SIMULATION RESULTS OF DESIGN OF AHB-LITE PROTOCOL

Figure 4. show single write and read operation which is taking place in AHB-Lite bus protocol. In the above simulated result the has been written by the signal Hw_data at the address Haddr when Hwrite signal is active high. The same data is been read by the system by the signal Hr_data at same address when the Hwrite signal is low.

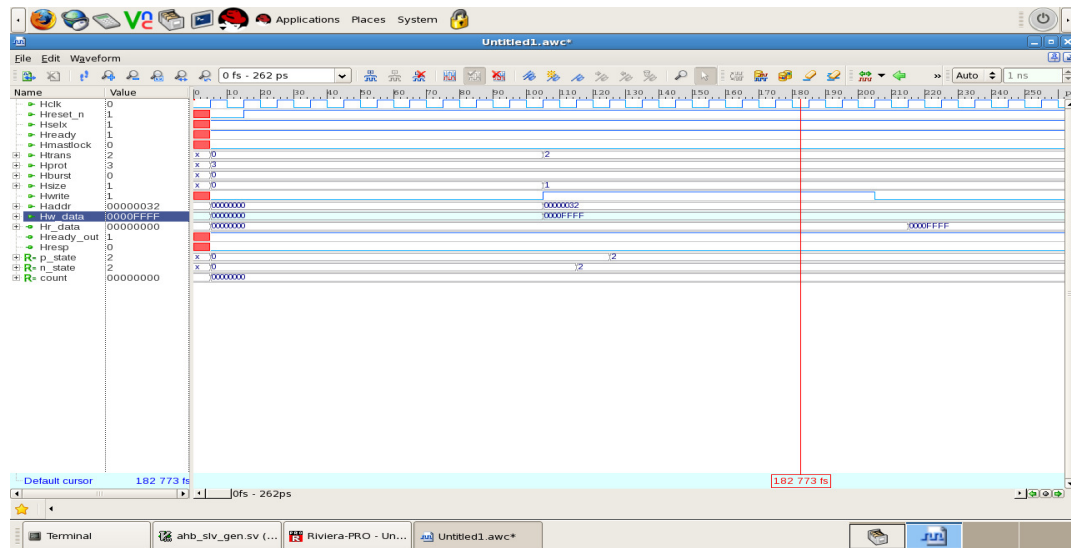


Figure 4. Single write and read operation

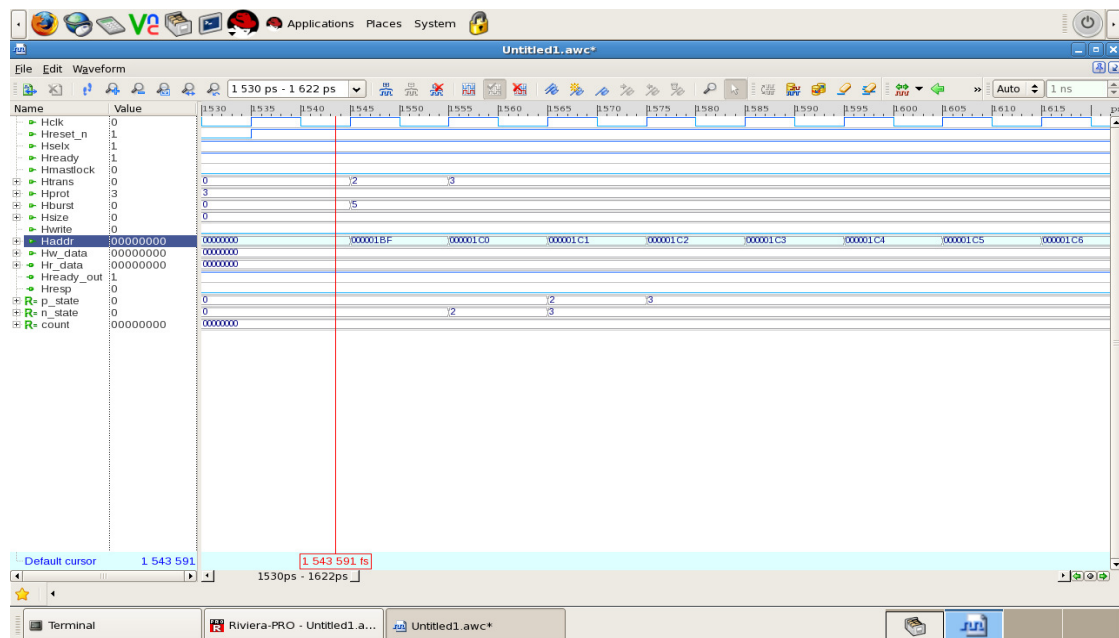


Figure 5. Read operation with unwritten

In Figure 5. Read operation with unwritten location is taking place i.e it is said to randomize operation. It is shows that the address is being added but Hw_data is 00000000 because this is taking place after the reset.

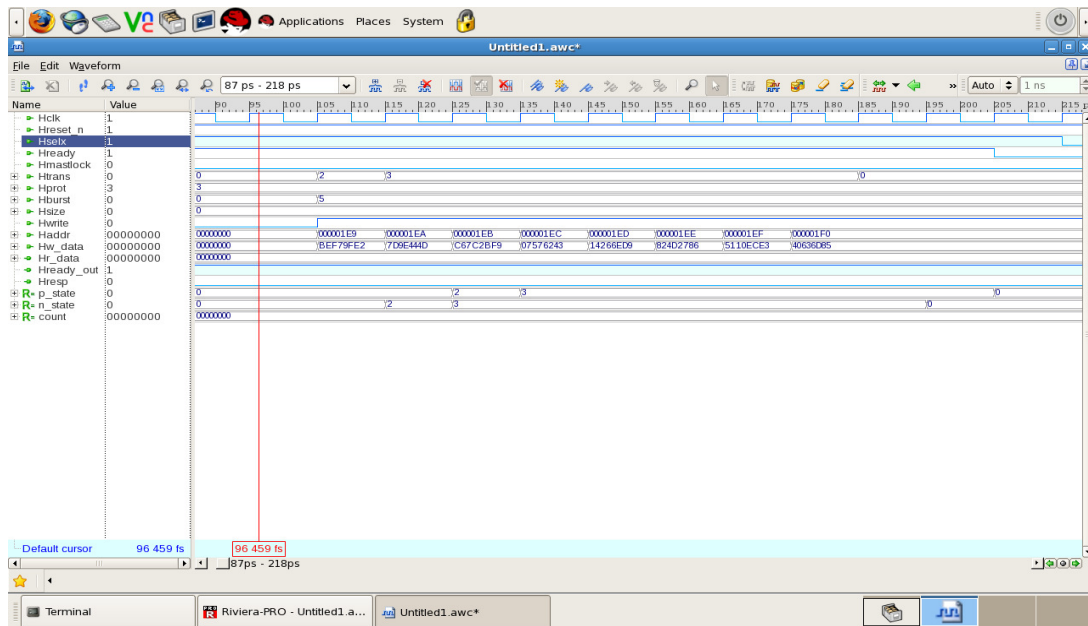


Figure 6. write_inc_8 operation

In Figure 6. it shown that inc_8 is taking place in Haddr of write. In Figure 7 inc_4 is taking place of write and read.

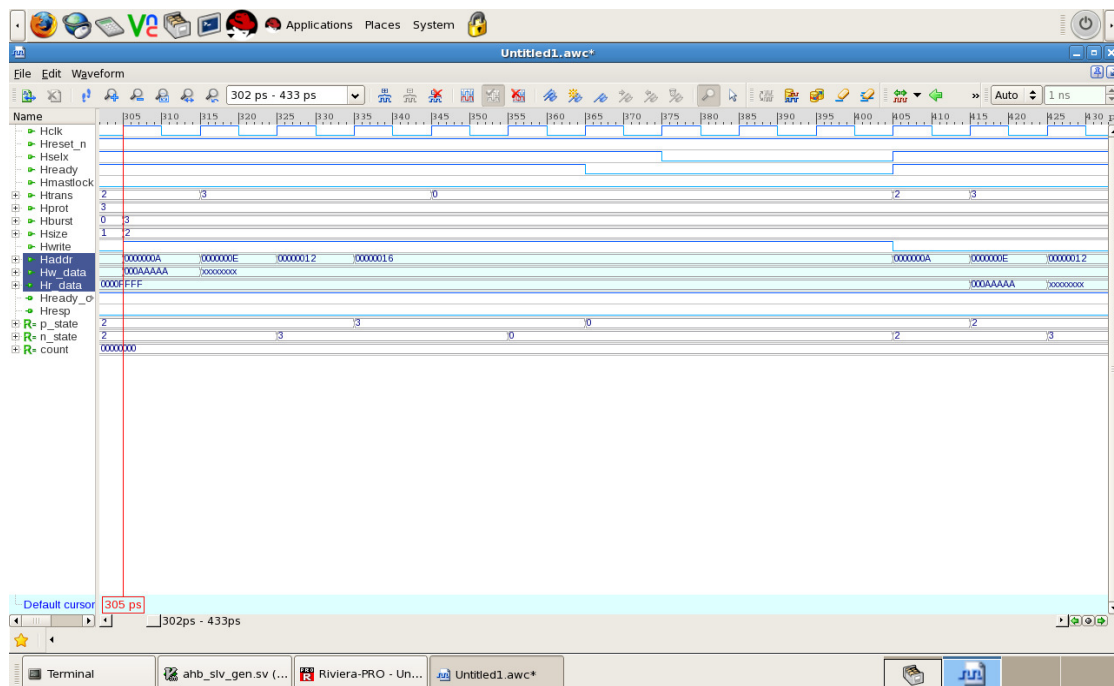


Figure 7. Write and read with inc_4

VIII. COVERAGE ANALYSIS

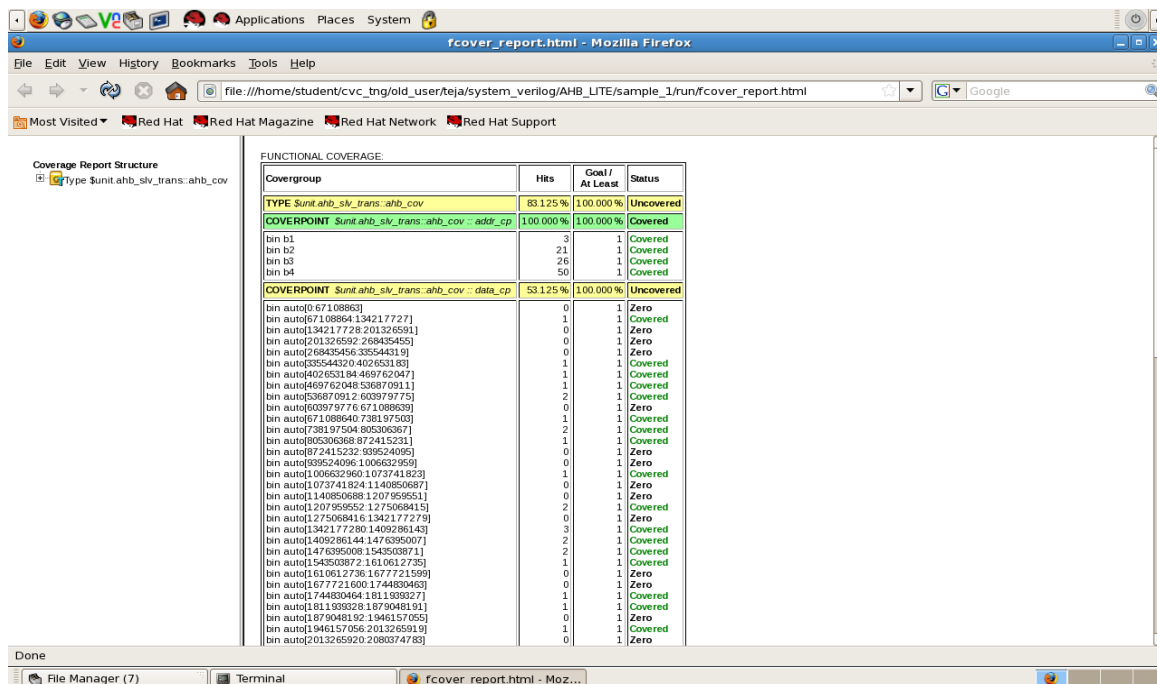


Figure 8. Coverage Analysis

The Coverage Report gives the details of the functional coverage when complete Analysis was done for the AHB-Lite and coverage report was generated as shown in Figure 8. It is found that the coverage is 100%.

IX. CONCLUSION

In the paper a general definition for AHB-LITE protocol which has high performance represents a significant advance in the capabilities of the ARM AMBA™ bus on-chip interconnect strategy, by providing a solution that reduces latencies and increases the bus bandwidth. AHB-Lite fully compatible with the current AHB specification. AHB-Lite increases the choice of architectures available to the AMBA bus-based designer, and is supported by a comprehensive range of products from ARM.

REFERENCES

- [1] ARM, "AMBA_3_AHB-Lite", available at <http://www.arm.com/>.
- [2] ARM, "AMBA Specification (Rev 2.0)", available at <http://www.arm.com>.
- [3] ARM, "AMBA AXI Protocol Specification", available at <http://www.arm.com>
- [4] Samir Palnitkar, "Verilog HDL: A Guide to Digital Design and Synthesis", Second Edition
- [5] Chris Spear, SystemVerilog for Verification, New York : Springer, 2006
- [6] <http://www.testbench.co.in>
- [7] http://www.doulos.com/knowhow/sysverilog/ovm/tutorial_0
- [8] <http://www.inno-logic.com/resourcesVMM.html>
- [9] Akhilesh Kumar and Richa Sinha, "Design and Verification Analysis of APB3 Protocol with Coverage" IJAET Journal, Vol. 1, Issue 5, pp. 310-317, Nov 2011.
- [10] http://en.wikipedia.org/wiki/Advanced_Microcontroller_Bus_Architecture
- [11] <http://books.google.co.in/books>
- [12] <http://www.asicguru.com>
- [13] <http://www.arm.com>
- [14] Bergeron, Janick. Writing testbenches: functional verification of HDL models. s.l.: Springer, 2003

Authors

Richa Sinha received B.E. Degree from RajaRamBapu Institute of Technology Shivaji University, Kolhapur, Maharashtra, India in 2007. Currently she is pursuing M. Tech project work under the guidance of Prof. Akhilesh Kumar in the Department of Electronics & Communication Engg, N. I. T., Jamshedpur. Her interest of field is ASIC Design & Verification.



Akhilesh Kumar received B.Tech degree from Bhagalpur University, Bihar, India in 1986 and M.Tech degree from Ranchi, Bihar, India in 1993. He has been working in teaching and research profession since 1989. He is now working as H.O.D. in Department of Electronics and Communication Engineering at N.I.T. Jamshedpur, Jharkhand, India. His interest of field of research is analog and digital circuit design in VLSI.



A. K. Sinha is presently Associate Professor and Head of the Department of Physics at S.G.G.S College, Patna Saheb, Bihar, India. She did her M.Sc (Physics) and Ph.D Degree from Magadh University, Bodh Gaya, Bihar in the month/year August 1981 and June 2003 respectively. Her fields of interest are Material Science, Semi-Conductor Devices, Electronic Spectra, Structure of Polyatomic Molecules, VLSI and its related fields.



IMPACT OF VOLTAGE REGULATORS IN UNBALANCED RADIAL DISTRIBUTION SYSTEMS USING PARTICLE SWARM OPTIMIZATION

Puthireddy Umapathi Reddy^{1*}, Sirigiri Sivanagaraju²,
Prabandhamkam Sangameswararaju³

¹Department of Electrical and Electronics Engineering, Sree Vidyanikethan Engineering
College, Tirupati, India.

²Department of Electrical and Electronics Engineering, Jawaharlal Nehru Technological
University College of Engineering Kakinada, Kakinada, India.

³Department of Electrical and Electronics Engineering, Sri Venkateswara University College
of Engineering, Tirupati, India.

ABSTRACT

In rural power systems, the Automatic Voltage Regulators (AVRs) help to reduce energy loss and to improve the power quality of electric utilities, compensating the voltage drops through distribution lines. This paper presents selection of optimal location and selection of tap setting for voltage regulators in Unbalanced Radial Distribution Systems (URDS). PSO is used for selecting the voltage regulator tap position in an unbalanced radial distribution system. An algorithm makes the initial selection, installation and tap position setting of the voltage regulators to provide a good voltage profile and to minimize power loss along the distribution network. The effectiveness of the proposed method is illustrated on a test system of 25 bus unbalanced radial distribution systems.

KEYWORDS: Unbalanced radial distribution systems, Voltage regulator placement, Loss minimization, Particle swarm optimization.

I. INTRODUCTION

This paper describes a new approach for modelling of automatic voltage regulator in the forward/backward sweep-based algorithms for unbalanced radial distribution systems [1], [2]. A voltage regulator is a device that keeps a predetermined voltage in a distribution network despite of the load variations within its rated power [3],[4]. Since it is the utilities' responsibility to keep the customer voltage within specified tolerances, voltage regulation is an important subject in electrical distribution engineering [5]. However, most equipment and appliances operate satisfactorily over some 'reasonable' range of voltages; hence, certain tolerances are allowed at the customers' end. Thus, it is common practice among utilities to stay within preferred voltage levels and ranges of variations for satisfactory operation of apparatus as set by various standards [6]. In distribution systems operation, shunt capacitor banks and feeder regulators are necessary for providing acceptable voltage profiles to all end-use customers and reducing power losses on large distribution systems [9]. A voltage regulator is equipped with controls and accessories for its tap to be adjusted automatically under load conditions. Moreover, it can be controlled by the installation of devices such as fixed and controlled capacitor banks, transformers with On Load Tap Changers (OLTCs), and Automatic Voltage Regulators (AVRs) [11], [12]. Loss reduction and improvement of voltage profile have been also studied by using OLTCs [13]. One of the most important devices to be utilized for the voltage

regulation is the AVRs which can be operated in manual or automatic mode. In the manual mode, the output voltage can be manually raised or lowered on the regulator's control board and it could be modelled as a constant ratio transformer in power flow algorithms [14]. In the automatic mode, the regulator control mechanism adjusts the taps to assure that the voltage being monitored is within certain range [16].

Optimal power flow analysis is used to determine the optimal tap position and the ON/OFF state of the capacitor banks. The same problem is solved by Vu et al. [7] using the loss equation as the objective function and voltage inequalities as constraints through the use of an artificial neural network. Safigianni and Salis [10] proposed the number and location of AVRs by using a sequential algorithm. In addition to this, the objective function is defined by using the AVR's investment and maintenance costs and also the cost of the total energy losses. Chiou et al. [15] initially attempted to solve the problem of voltage regulator by changing the tap positions at the substation and later solved the capacitor problem. J. Mendoza et al. [17] developed a method for optimal location of AVRs in radial distribution networks by using simple genetic algorithms. However, there are only a few publications that have treated the complex problem of the optimal location of the AVRs in distribution systems, despite the fact that the benefits of including AVR devices. The automatic voltage regulators (AVRs) are included into the sweep based methods and tested by using two distribution test systems. The proposed method deals with the placing of voltage regulator and tap position of regulators for power loss reduction and voltage profile improvement [18]. Daniela Proto, Pietro Varilone et al [19] explained about Voltage Regulators and Capacitor Placement in Three-phase Distribution Systems with Non-linear and Unbalanced Loads. Multiobjective Location of Automatic Voltage regulators in a radial Distribution Network Using a Micro Genetic Algorithm is given in [20]. Optimal Distribution Voltage Control and losses minimization suitable methods are proposed in [21], [22]. Integrated volt/VAr control in distribution systems is illustrated in [23],[24]. Impact of Distribution Generation on voltage Levels in Radial Distribution Systems using Voltage regulators/controllers in [26],[27] are proposed to improve voltage profile.

This paper, explains mathematical model, Algorithm for finding the tap settings of a voltage regulator, Implementation of PSO and results and discussions. The branch that has the highest voltage drop is picked as the best location for the voltage regulator placement. PSO is used to find the selection of tap position of the voltage regulator. to obtain the tap position of the voltage regulators that maintain the voltages within the limits of the unbalanced radial distribution systems so as to minimize an objective function, which consists of power loss.

II. MATHEMATICAL FORMULATION

In this paper in order to maintain the voltage profile and to reduce the power losses, these voltage regulators are installed in the distribution system. The optimization problem has been presented into two sub problems: Locating the AVRs on the network and the selection of the tap position of AVRs.

2.1 Optimal location of Automatic Voltage Regulators (AVR)

The optimal location of voltage regulator (AVR) is defined as function of two objectives, one representing power loss reduction and the other one representing the voltage deviations but both are essential to secure the power supply. It is difficult to formulate the problem in terms of cost incidence of these objectives over the system of operation because even when the cost incidence of power losses is clear it is not the same for keeping the voltage values at the buses close to the rated value.

The objective function to be minimized is

$$\text{Minimize } f = \sum_{j=1}^{nb} P'_{\text{loss } j} \quad (1)$$

Where, $P'_{\text{loss } j}$ is the active power loss in the j^{th} branch after voltage regulator placement.

'nb' is the number of branches in the system.

2.2 Tap Position Selection

In general, voltage regulator position at bus 'q' can be calculated as

$$V_q^{ph'} = V_q^{ph} \pm \text{Tap}^{ph} \times V_{\text{rated}}^{ph} \quad (2)$$

Tap position (tap) can be calculated by comparing voltage obtained before VR installation with the lower and upper limits of voltage.

'+' for boosting of voltage

'-' for bucking of voltage

The bus voltages are computed by load flow analysis for every change in tap setting of voltage regulators, till all bus voltages are within the specified limits.

III. ALGORITHM FOR FINDING THE TAP SETTINGS OF A REGULATOR

Step 1 : Read the system and regulator data

Step 2 : Calculate the branch current in which regulator is inserted from the backward sweep.

Step 3 : Find the CT ratio for three phases as

$$CT_{ph} = \frac{CT_P^{ph}}{CT_S^{ph}} \quad \text{Whereas } CT_S^{ph} = 5 \text{ Amps}, \quad (3)$$

Step 4 : Convert the R and X values from volts to ohms as

$$(R - jX)_{ohms}^{ph} = \frac{(R_{\text{Setting}} - jX_{\text{Setting}})^{ph} \text{ volts}}{CT_S \text{ of current }^{ph}} \quad (4)$$

Step 5: Calculating current in the compensator

$$I_{comp}^{ph} = \frac{(\text{current in the branch})^{ph}}{CT \text{ ratio }^{ph}} \quad (5)$$

Step 6: Calculate the input voltage to the compensator as

$$V_{reg}^{ph} = \frac{(\text{Voltage at the sending end of the branch})^{ph}}{PT \text{ ratio }^{ph}} \quad (6)$$

Step 7: Voltage drop in the compensator circuit is

$$V_{drop}^{ph} = (R + jX)_{ohms}^{ph} I_{comp}^{ph} \quad (7)$$

Step 8: Voltage across the voltage relays in three phases

$$V_R^{ph} = V_{reg}^{ph} - V_{drop}^{ph} \quad (8)$$

Step 9: Finding the tapping of the regulator

$$\text{Tap}^{ph} = \frac{(\text{lower limit of the voltage limit})^{ph} - V_R^{ph}}{(\text{change in voltages for a step change of the regulator})^{ph}} \quad (9)$$

Step 10: Voltage output of the regulator

$$V_{ro}^{ph} = (\text{voltage of the sending end of the branch})^{ph} \pm \text{Tap}^{ph} \times (0.00625, \quad (10)$$

'+' For raise

'-' For lower

Step 11: Stop

IV. IMPLEMENTATION OF PSO

In this section, the optimal voltage regulator tap setting at candidate node of the unbalanced radial distribution system is selected using PSO.

4.1 Initialization of PSO Parameters

The control parameters such as lower and upper bounds of node voltage and tap setting of voltage regulators are selected as initialize parameters. Randomly generate an initial swarm (array) of particles with random positions and velocities.

4.2 Evaluation of Fitness Function

The fitness function should be capable of reflecting the objective and directing the search towards optimal solution. Since the PSO proceeds in the direction of evolving best-fit particles and the fitness value is the only information available to the PSO, the performance of the algorithm is highly sensitive to the fitness values. For each particle or swarm, the voltage regulators are placed at the sensitive nodes and run the load flow to calculate the losses, net saving using Eqn.(1) and these net saving becomes the fitness function of the PSO (as saving are maximized).

4.3 Optimal Solution

Optimal solution (the best position and the corresponding fitness value) to the target problem. Information of the best position includes the optimal location and number of voltage regulators, and the corresponding tap setting value represents the maximizing the total saving of the system. Accordingly, the optimal location and number of voltage regulators with tap setting at each node can be determined.

This modification can be represented by the concept of velocity (modified value for the current positions). Velocity of each particle can be modified by the following equation

$$V_i^{k+1} = WV_i^k + C_1 rand_1 \times [Pbest_i - X_i^k] + C_2 rand_2 \times [Gbest - X_i^k] \quad (11)$$

where,

V_i^k : Velocity of particle i at iteration k,

V_i^{k+1} : Modified velocity of particle i at iteration k+1,

W : Inertia weight,

C_1, C_2 : Acceleration Constants,

$rand_1, rand_2$: Two random numbers

X_i^k : Current position of particle i at iteration k,

$Pbest_i$: Pbest of particle i,

$Gbest$: Gbest of the group.

In the equation (4),

The term $rand_1 \times (Pbest_i - X_i^k)$ is called particle memory influence

The term $rand_2 \times (Gbest - X_i^k)$ is called swarm influence.

The $rand_1, rand_2$ are the two random numbers with uniform distribution with range of { 0.0 to 1.0 }

W is the inertia weight which shows the effect of previous velocity vector on the new vector. Suitable selection of inertia weight W provides a balance between global and local exploration, thus requiring less iteration on average to find optimal solution. A larger inertia weight 'W' facilitates global exploration, while smaller inertia weight 'W' tends to facilitates local exploration to fine tune.

The following inertia weight is usually utilized in equation (11):

$$W = W_{\max} - \frac{W_{\max} - W_{\min}}{iter_{\max}} \times iter \quad (12)$$

Where,

W_{\max} : Initial value of the Inertia weight,

W_{\min} : Final value of the Inertia weight,

$iter_{\max}$: Maximum iteration number,

$iter$: current iteration number.

Accordingly, the optimal types and sizes of voltage regulators to be placed at each compensation node can be determined.

4.4 Algorithm for Optimal Location using PLI and Tap setting of VR using PSO

The detailed algorithm is to determine optimal location along with tap setting of voltage regulator is given below.

- Step 1: Read system data such as line data and load data of distribution system.
- Step 2: Initialize the PSO parameters such as Number of Agents (M), Number of Particles (N), Number of Iterations (K_{max}), Initial value of Inertia weight (W_{max}), Final value of Inertia weight (W_{min}), Acceleration Constants (C_1 & C_2).
- Step 3: Initialize the parameters as explained in section 4.(i)
- Step 4: Obtain the optimal location of VR by using PLI (Power Loss Index) as input.
- Step 5: Initialize the swarm by assigning a random position and velocity in the problem hyperspace to each particle, where each particle is a solution of tap setting of VR.
- Step 6: Run the load flow and compute the fitness value of each particle using equation (11).
- Step 7: Compare the present fitness value of i^{th} particle with its historical best fitness value. If the present value is better than Pbest update the Pbest, else retain Pbest as same.
- Step 8: Find the Gbest value from the obtained Pbest values.
- Step 9: Update the particle positions & velocity using eqns. (11) & (12).
- Step 10: Apply boundary conditions to the particles
- Step 11: Execute steps 6-10 in a loop for maximum number of iterations (K_{max}).
- Step 12: Stop the execution and display the Gbest values as the final result for optimal tap setting of voltage regulator.

V. RESULTS AND DISCUSSION

The performance of the proposed method is evaluated for test system of 25 bus URDS for voltage regulator placement to find placing and tap settings of the voltage regulator. For the positioning of voltage regulators, the upper and lower bounds of voltage are taken as $\pm 5\%$ of base value. Proper allocation of VR gives minimum losses in URDS and improves performance of the system. The real and reactive power losses of given test system is controlled by controlling voltage regulator size and location. The PSO Parameter values for voltage regulator placement: Number of Particles (N) =20, Number of Iterations (K_{max}) =100, Initial value of the inertia weight (W_{max}) =0.9 ,Final value of the inertia weight (W_{min}) =0.4, Acceleration constants, (C_1 & C_2)=4. The proposed method is illustrated with test system consisting of 25 bus URDS.

5.1 Case Study

Power loss indices for 25 bus URDS is shown in the Figure.1. From the Figure.1 it can be concluded that the power loss index above 0.6 is the minimum voltage point to locate voltage regulator. The proposed algorithm is tested on 25 bus URDS as shown in Figure 2. The line and the load data of this system is given in [18]. The tap settings of the regulator are obtained with PSO algorithm. The single line diagram of 25 bus URDS after voltage regulator placement is shown in Figure.2

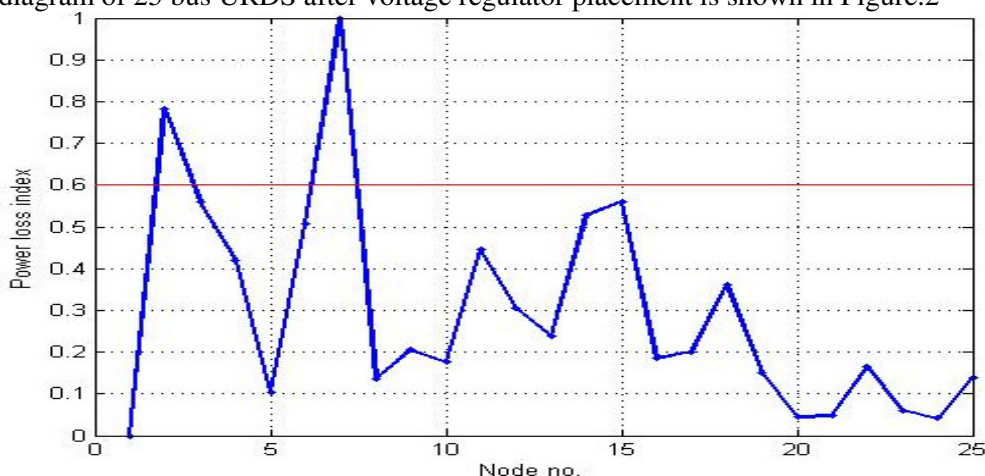


Figure.1 Power loss indices for 25 bus URDS

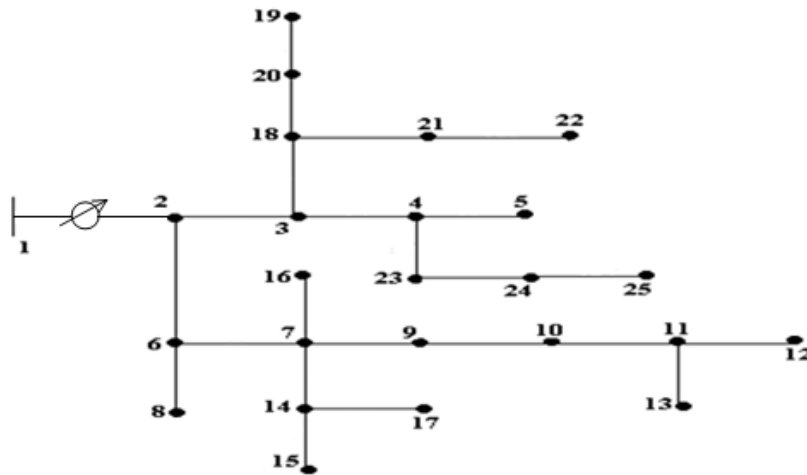


Figure. 2 Single line diagram of 25 bus URDS after voltage regulator placement.

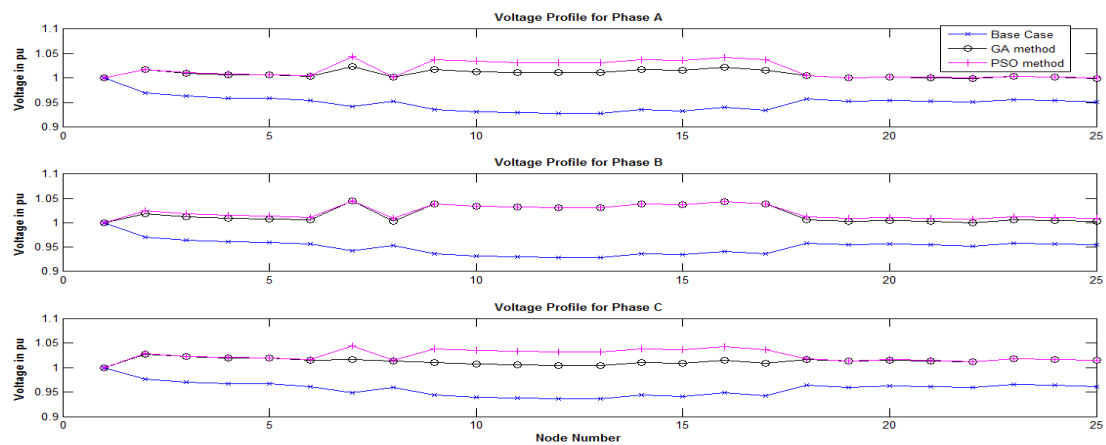


Figure.3 Voltage profit

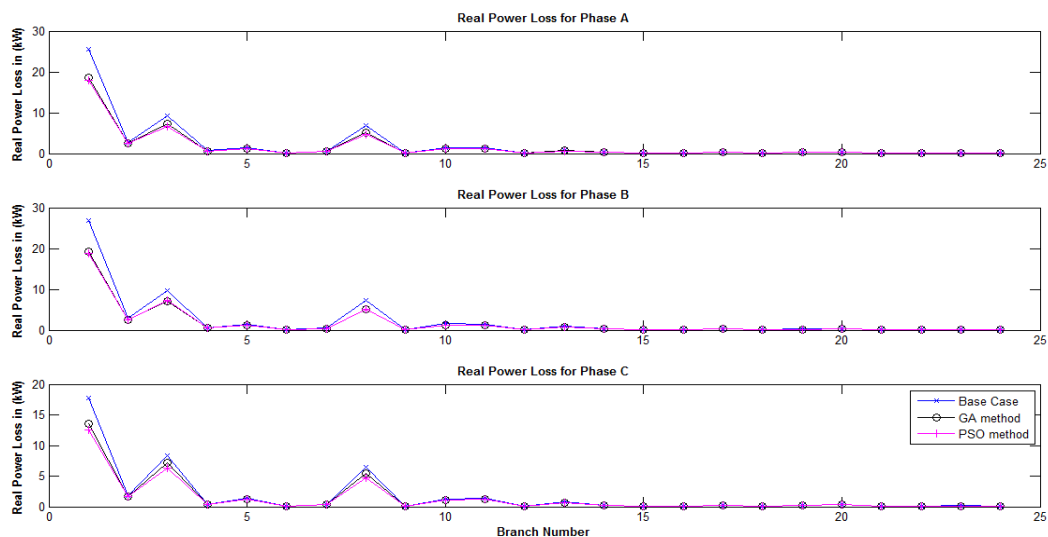


Figure 4. Real power loss

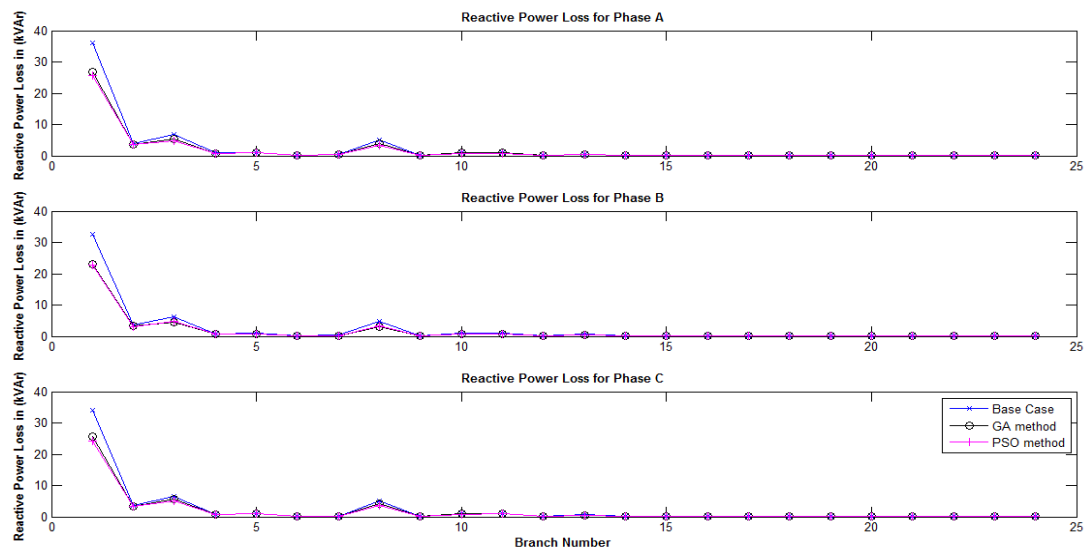


Figure.5 Reactive power losses

Table: 1 Summary of test results of 25 bus URDS with voltage regulator placement.

Description		Without Regulators			With Regulators					
					Existing method[18]			Proposed method		
	Node.N o.	Phase A	Phase B	Phase C	Phase A	Phase B	Phase C	Phase A	Phase B	Phase C
Tap-setting of reg. with node no.	2 7				7 5	7 8	8 2	7 8	8 7	8 6
Min. voltage (p.u)		0.9284	0.9284	0.9366	0.9995	1.0000	1.0000	1.0000	1.0000	1.0000
Max. Voltage Regulation(p.u)		0.0726	0.0716	0.0634	0.0005	0	0	0	0	0
Improvement in Maximum Voltage Regulation (%)		---	---	---	99.3113	100	100	100	100	100
Total active Power demand(kW)		1126.1164	1138.744	1125.1621	1114.2061	1124.6598	1117.7285	1112.4587	1124.2917	1114.9545
Total active power loss(kW)		52.8148	55.4428	41.8615	40.9061	41.3598	34.4285	39.1587	40.9917	31.6545
Improvement in active power loss reduction (%)		---	---	---	22.5480	25.4009	17.7562	25.84657	26.0649	24.3828
Total reactive Power demand(kVar)		850.3216	854.2954	855.6916	837.1748	839.9838	844.4912	835.1034	839.9680	841.4420
Total reactive power loss (kVar)		58.3198	53.2942	55.6909	45.1748	38.9838	44.4912	43.1034	38.9680	41.4420
Improvement in reactive power loss reduction (%)		---	---	---	22.5395	26.8517	20.1105	26.0913	26.8813	25.5857
Total Feeder Demand (kVA)		1411.0935	1423.5724	1413.5763	1393.6707	1403.7209	1400.8863	1391.0291	1403.4166	1396.8351
Released feeder capacity (kVA)		---	---	---	17.4228	19.8515	12.69	20.0645	20.156	16.7412
Total system power loss (kW.)	Best	---			116.6944			111.8049		
	worst				150.1191			118.5229		
	Avg				117.5996			112.0087		
Time of Execution(s)					73.7340			66.20300		

From the Figure.2 it can be concluded that 1st branch having more drop than others. Therefore voltage regulator should be placed in this branch. It is boosted the total network voltage intern power losses are minimized.

Voltage profit values, active and reactive power loss and summary of test results of 25 bus URDS for voltage regulator placement are given in figure.3, figure.4 and table.1 respectively. From table 1, it is observed that the minimum voltages in phases A, B and C are improved from 0.9284, 0.9284 and 0.9366 p.u (without Regulators) to 1.0, 1.0 and 1.0 p.u (with Regulators) respectively and the active power loss in phases of A, B and C is reduced from 52.82, 55.44 and 41.86 kW to 39.16, 40.99 and 31.65 kW respectively. Hence, there is an improvement in the minimum voltage and reduction in

active power loss when compared with the before voltage regulator placement and after voltage regulator placement. The total active power loss Vs generation number of 25 bus URDS is shown in Figure 6.

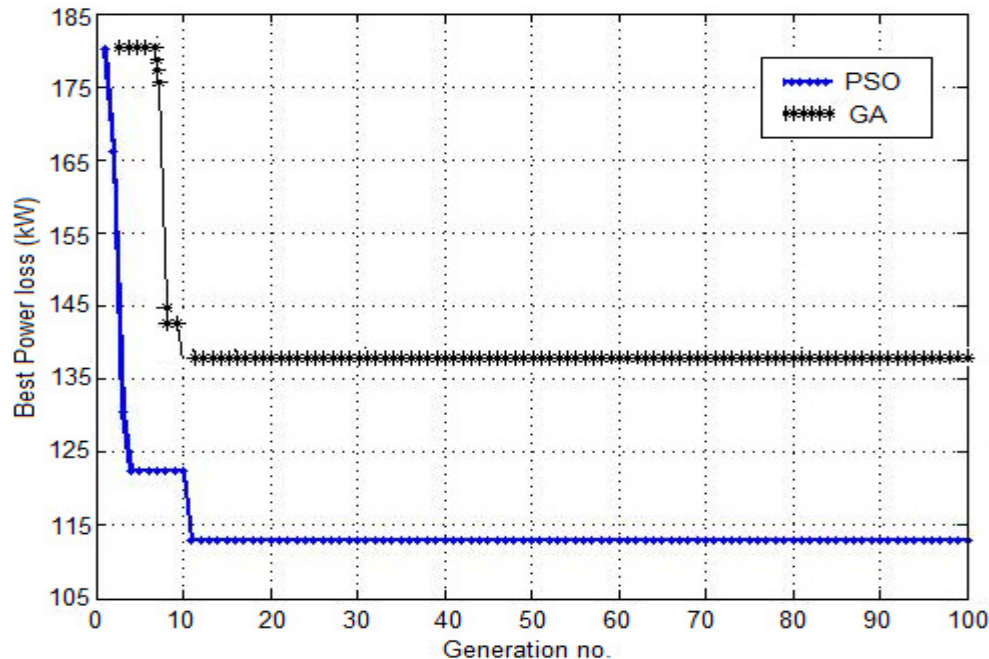


Figure. 6 Total Active Power loss Vs Generation number of 25 bus URDS

VI. CONCLUSIONS

This paper presents a simple method to determine optimal allocation and tap setting of voltage regulators in Unbalanced Radial Distribution Systems through voltage drop analysis and PSO algorithm respectively. The effectiveness of the PSO has been demonstrated and tested. The proposed PSO based methodology was applied to 25 bus URDS. The obtained solution has succeeded in reducing total active power losses 25.43% in 25 bus URDS. From the test results, it can be said that the proposed model is valid and reliable, and the performance of the algorithms is not significantly affected from the inclusion of regulator modelling. The power loss per phase of unbalanced distribution system can be reduced by proper placement of voltage regulator. In addition to power loss reduction, the voltage profile also improved by the proposed method. The time of execution is reduced from 73.7 to 66.20 seconds for the same configuration system.

REFERENCES

- [1] R.R. Shouts, M.S. Chen, and L. Schwobel, "Simplified feeder modelling for load flow calculations", IEEE Transactions Power Systems, vol.2, pp.168-174, 1987.
- [2] T.-H. Chen, M.-S. Chen, K.J. Hwang, P. Kotas, and E. A. Chebli, "Distribution system power flow analysis – A rigid approach", IEEE Transactions on Power Delivery, vol. 6, pp. 1146–1152, July 1991.
- [3] D. Rajicic, R. Ackovski, and R. Taleski, "Voltage correction power flow", IEEE Transactions on Power Delivery, vol. 9, pp. 1056–1062, Apr. 1994.
- [4] S.K. Chang, G. Marks, and K. Kato, "Optimal real time voltage control", IEEE Transactions Power Systems, vol. 5, no. 3, pp. 750–758, Aug. 1990.
- [5] C. J. Bridenbaugh, D. A. DiMascio, and R. D'Aquila, "Voltage control improvement through capacitor and transformer tap optimization", IEEE Transactions Power Systems, vol. 7, no. 1, pp. 222– 226, Feb. 1992.
- [6] C. S. Cheng and D. Shirmohammadi, "A Three Phase Power Flow Method for Real Time Distribution System Analysis", IEEE Transactions on Power Systems, vol. 10, no. 2 pp 671- 679, May 1995.
- [7] H. Vu, P. Pruvot, C. Launay, and Y. Harmand, "An improved voltage control on large-scale power systems", IEEE Transactions Power Systems, vol. 11, no. 3, pp. 1295–1303, Aug. 1996.

- [8] Z. Gu and D. T. Rizy, "Neural network for combined control of capacitor banks and voltage regulators in distribution systems", *IEEE Transactions Power Delivery*, vol. 11, no. 4, pp. 1921–1928, Oct. 1996.
- [9] M. M. A Salama, N. Manojlovic, V. H. Quintana, and A. Y. Chikhani, "Real-Time Optimal Reactive Power Control for Distribution Networks", *International Journal of Electrical Power & Energy Systems*, vol. 18, no. 3, pp. 185–193, 1996.
- [10] A. Safigianni and G. Salis, "Optimal voltage regulator placement in radial distribution network," *IEEE Trans. on Power Systems*, vol. 15, no. 2, pp. 879–886, May 2000.
- [11] M. A. S. Masoum, A. Jafarian, M. Ladjevardi, E. F. Fuchs, and W. N Grady, "Fuzzy approach for optimal Placement and sizing of capacitor banks in the presence of harmonic", *IEEE Transactions Power Delivery*, vol. 16, no. 2, pp. 822–829, Apr. 2004.
- [12] B. Alencar de Souza, H. do Nascimento Alves, and H. A. Ferreira, "Micro genetic algorithms and fuzzy logic applied to the optimal placement of capacitor banks in distribution networks," *IEEE Transactions Power Systems*, vol. 19, no. 2, pp. 942–947, May 2004.
- [13] B. Milosevic and M. Begovic, "Capacitor placement for conservative voltage reduction on distribution feeders," *IEEE Transactions Power Delivery*, vol. 19, no. 3, pp. 1360–1367, July 2004.
- [14] M. A. S. Masoum, M. Ladjevardi, A. Jafarian, and E. Fuchs, "Optimal placement, replacement and sizing of voltage regulators in distorted distribution networks by genetic algorithms", *IEEE Transactions Power Delivery*, vol. 19, no. 4, pp. 1794–1801, Oct. 2004.
- [15] J. Chiou, C. Chang, and C. Su, "Ant direction hybrid differential evolution for solving large capacitor placement problems", *IEEE Transactions Power Systems*, vol. 19, no. 4, pp. 1794–1800, Nov. 2004.
- [16] A. Augugliaro, L. Dusonchet, S. Favazza, and E. Riva, "Voltage regulation and power losses minimization in automated distribution networks by an evolutionary multiobjective approach," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1516–1527, Aug. 2004.
- [17] J. Mendoza et al "optimal location of voltage regulators in radial distribution networks using genetic algorithms," in *Proceedings 15th power systems computation conference*, Belgium, August 2005.
- [18] J.B.V. Subramnyam, "Optimal capacitor placement in unbalanced radial distribution networks *Journal of Theoretical and Applied Information Technology* vol:6,N0:1,pp106-115. 2009.
- [19] Daniela Proto, Pietro Varilone, " Voltage Regulators and Capacitor Placement in Three-phase Distribution Systems with Non-linear and Unbalanced Loads" *International Journal of Emerging Electric Power Systems*, Vol. 7, No.4, Nov. 2010.
- [20] J.E. Mendoza, D.A. Morales, R.A.Lopez, E.A.Lopez, "Multiobjective Location of Automatic Voltage regulators in a radial Distribution Network Using aMicro Genetic Algorithm" *IEEE Transactions Power Systems*, vol. 22, no. 1, pp. 404–412, Feb. 2007.
- [21] T.Senju, Y.Miyazato, A. Yona, N.Urasaki, T.Funabashi, " Optimal Distribution Voltage Control and Coordination With Distributed Generation" *IEEE Transactions Power Delivery*, vol. 23, no. 2, pp. 1236–1242, April 2008.
- [22] H.A Attia, " Optimal voltage profile control and losses minimization of radial distribution feeder" *Power System Conference, (MEPCON 2008)*, pp 453-458, March 2008.
- [23] P.V.V.RamaRao, S.Sivanagaraju, "Voltage Regulator Placement In Radial distribution Network Using Plant Growth Simulation Algorithm" *International Journal of Engineering, Science and Technology*, Vol. 2, No. 6, pp. 207-217, 2010.
- [24] V. Borozan, M.E.Baran, D. Novosel, " Integrated volt/VAr control in distribution systems" *IEEE Power Engineering Society Winter Meeting*, vol. 3, pp. 1485–1490, Feb 2010.
- [25] B.A. De Souza, A.M.F de Almeida, " Multi objective Optimization and Fuzzy Logic Applied to Planning of the Volt/ Var Problem in Distributions Systems" *IEEE Transactions Power Systems*, vol. 25, no. 3, pp. 1274–1281, Aug. 2010.
- [26] Srikanth Apparaju, Sri Chandan K "impact of Distribution Generation on voltage Levels in Radial Distribution Systems" *International Journal of Engineering Research and Applications* Vol. 1, Issue 2, pp.277-281, 2010.
- [27] Jianzhong Tong; Souder, D.W. Piong, C. Mingye Zhang, Qinglai Guo, Hongbin Sun, Boming Zhang, " Voltage control practices and tools used for system voltage control of PJM" *IEEE power and Energy Society General Meeting*, pp.1-5, July 2011.

Authors

P. Umapathi Reddy: He Received B.E from Andra University and M.Tech.,(Electrical Power Systems) from Jawaharlal Nehru Technological University, Anantapur, India in 1998 and 2004 respectively, Now he is pursuing Ph.D. degree. Currently he is with Department of Electrical and Electronics Engineering, Sree Vidyanikethan Engineering College, Tirupati, India. His research interest includes Power distribution Systems and Power System operation



and control. He is Life Member of Indian Society for Technical Education.

S. Sivanaga Raju: He received B.E from Andra University and M.Tech.degree in 2000 from IIT, Kharagpur and did his Ph.D from Jawaharlal Nehru Technological University, Anantapur, India in 2004. He is presently working as Associate professor in J.N.T.U.College of Engineering Kakinada,(Autonomous) Kakinada, Andrapradesh, India. He received two national awards (Pandit Madan Mohan Malaviya memorial Prize and best paper prize award from the Institute of Engineers (India) for the year 2003-04. He is referee for IEEE journals. He has around 75 National and International journals in his credit. His research interest includes Power distribution Automation and Power System operation and control.



P. Sangameswara Raju He is presently working as professor in S.V.U.College Engineering, Tirupati. Obtained his diploma and B.Tech in electrical Engineering, M.Tech in power system operation and control and Ph.D in S.V.University, Tirupati. His areas of interest are power system operation, planning and application of fuzzy logic to power system, application of power system like non-linear controllers.



STUDY ON PERFORMANCE OF CHEMICALLY STABILIZED EXPANSIVE SOIL

P. VenkaraMuthyalu, K. Ramu and G.V.R. Prasada Raju

Department of Civil Engg., University College of Engineering, JNTUK, Kakinada, India

ABSTRACT

Expansive soils, such as black cotton soils, are basically susceptible to detrimental volumetric changes, with changes in moisture. This behaviour of soil is attributed to the presence of mineral montmorillonite, which has an expanding lattice. Understanding the behaviour of expansive soil and adopting the appropriate control measures have been great task for the geotechnical engineers. Extensive research is going on to find the solutions to black cotton soils. There have been many methods available to controlling the expansive nature of the soils. Treating the expansive soil with electrolytes is one of the techniques to improve the behaviour of the expansive ground. Hence, in the present work, experimentation is carried-out to investigate the influence of electrolytes i.e., potassium chloride, calcium chloride and ferric chloride on the properties of expansive soil.

KEYWORDS: *Expansive soil, Calcium Chloride, Potassium Chloride, Ferric Chloride*

I. INTRODUCTION

Expansive soil is one among the problematic soils that has a high potential for shrinking or swelling due to change of moisture content. Expansive soils can be found on almost all the continents on the Earth. Destructive results caused by this type of soils have been reported in many countries. In India, large tracts are covered by expansive soils known as black cotton soils. The major area of their occurrence is the south Vindhya range covering almost the entire Deccan Plateau. These soils cover an area of about 200,000 square miles and thus form about 20% of the total area of India. The primary problem that arises with regard to expansive soils is that deformations are significantly greater than the elastic deformations and they cannot be predicted by the classical elastic or plastic theory. Movement is usually in an uneven pattern and of such a magnitude to cause extensive damage to the structures resting on them.

Proper remedial measures are to be adopted to modify the soil or to reduce its detrimental effects if expansive soils are identified in a project. The remedial measures can be different for planning and designing stages and post construction stages. Many stabilization techniques are in practice for improving the expansive soils in which the characteristics of the soils are altered or the problematic soils are removed and replaced which can be used alone or in conjunction with specific design alternatives. Additives such as lime, cement, calcium chloride, rice husk, fly ash etc. are also used to alter the characteristics of the expansive soils. The characteristics that are of concern to the design engineers are permeability, compressibility and durability. The effect of the additives and the optimum amount of additives to be used are dependent mainly on the mineralogical composition of the soils. The paper focuses about the various stabilization techniques that are in practice for improving the expansive soil for reducing its swelling potential and the limitations of the method of stabilization there on.

Modification of BC soil by chemical admixture is a common method for stabilizing the swell-shrink tendency of expansive soil [5]. Advantages of chemical stabilization are that they reduce the swell-shrink tendency of expansive soils and also render the soils less plastic. Among the chemical

stabilization methods for expansive soils, lime stabilization is mostly adopted for improving the swell-shrink characteristics of expansive soils. The reaction between lime and clay in the presence of water can be divided into two distinct processes [20]. The use of calcium chloride in place of lime, as calcium chloride is more easily made into calcium charged supernatant than lime [40]. The electrolytes like potassium chloride, calcium chloride and ferric chloride can be effectively used in place of the conventionally used lime, because of their ready dissolvability in water and supply of adequate cations for ready cation exchange ([55],[56],[42]).

Calcium chloride is known to be more easily made into calcium charged supernatant than lime and helps in ready cation exchange reactions [44]. The CaCl_2 might be effective in soils with expanding lattice clays [33]. The stabilization of the in-situ soil using KOH solution was made and revealed that the properties of black cotton soils in place can be altered by treating them with aqueous solution of KOH [27]. The laboratory tests reveal that the swelling characteristics of expansive soils can be improved by means of flooding at a given site with proper choice of electrolyte solution more so using chloride of divalent or multivalent cations [19]. The influence of CaCl_2 and KOH on strength and consolidation characteristics of black cotton soil is studied [55] and found an increase in the strength and reduction in the settlement and swelling. 5% FeCl_3 solution to treat the caustic soda contaminated ground of an industrial building in Bangalore [55]. In this work an attempt is made to study the effect of electrolytes like KCl, CaCl_2 and FeCl_3 on the properties of expansive soil.

The bibliography on stabilization of soil and calcium chloride giving its wide use in highways [58]. [30],[18], [53] has stated that CaCl_2 enjoyed its wide use as dust palliative and frost control of subgrade soil.

When lime stabilization is intended to modify the in-situ expansive soil bed it is commonly applied in the form of lime piles ([24],[6],[23],[7],[1],[10],[65],[18],[51]) or lime slurry pressure injection (LPSI) ([66],[63],[36],[58],[26],[9],[3],[59]).

Numerous investigators, ([20], [34], [64], [43], [15], [41], [35], [45], [29], [37], [45], [4], [22], [2], [31], [39], [32]), have studied the influence of lime, cement, lime-cement, lime-flyash, lime-rice-husk-ash and cement-flyash mixes on soil properties, mostly focusing on the strength aspects to study their suitability for road bases and subbases. As lime and cement are binding materials, the strength of soil-additive mixtures increases provided the soil is reactive with them. However, for large-scale field use, the problems of soil pulverization and mixing of additives with soil have been reported by several investigators ([20],[58],[9],[5],[44]).

It is an established fact that, whenever a new material or a technique is introduced in the pavement construction, it becomes necessary to experiment it for its validity by constructing a test track, where the loading, traffic and other likely field conditions are simulated. Several test track studies ([38],[49],[54],[50],[12],[25],[8],[14],[17],[52]), have been carried out in many countries to characterize the pavement materials and to assess the effectiveness of remedial techniques developed to deal with the problematic condition like freeze-thaw, expansive soil and other soft ground problems.

Recent studies ([60],[28]), indicated that CaCl_2 could be an effective alternative to conventional lime used due to its ready dissolvability in water and to supply adequate calcium ions for exchange reactions. [13] Studied the use of KCl to modify heavy clay in the laboratory and revealed that from engineering point of view, the use of KCl as a stabilizer appears potentially promising in locations where it is readily and cheaply available. In the present work, the efficiency of Potassium Chloride (KCl), Calcium Chloride (CaCl_2) and Ferric Chloride (FeCl_3), as stabilizing agents, was extensively studied in the laboratory for improving the properties of expansive soil.

The experiences of various researchers in the field as well as laboratory chemical stabilization have been presented briefly in the above section. Experimental study methodologies for laboratory are presented in the following section.

II. EXPERIMENTAL STUDY

2.1. Soil

The black cotton soil was collected from Morampalem, a village nearer to Amalapuram of East Godavari District in Andhra Pradesh in India. The physical properties of the soil are given in Table 1.

Table.1: Physical Properties of Expansive Soil

Property					
Grain Size Distribution	Sand (%)	Silt (%)	Clay (%)		
	2	22	76		
Atterberg Limits	Liquid limit (%)	Plastic limit (%)	Plasticity Index	Shrinkage Limit (%)	
	85	39	46	12	
Classification	CH	Specific Gravity	2.68	Free Swell Index	140 %
Compaction properties	Maximum Dry Density (g/cc)	Optimum moisture Content (%)	Soaked CBR of sample prepared at MDD & OMC	2	
	1.42	26.89			
Permeability of the sample prepared at OMC & MDD	1.89×10^{-7} cm/sec	Shear Strength Parameters of the sample prepared at OMC & MDD	Cohesion (C) (kg/cm ²)	Angle of internal friction (ϕ)	
			0.56	2^0	

2.2. Chemicals

Three chemicals of commercial grade, KCl, CaCl₂ and FeCl₃ are taken in this study. The quantity of the chemical added to the expansive soil was varied from 0 to 1.5% by dry weight of soil.

2.3. Test Program

Electrolytes like KCl, CaCl₂ and FeCl₃ are mixed in different proportions to the expansive soil and the physical properties like liquid limit, plastic limit, shrinkage limit and DFS of the stabilized expansive soil are determined to study the influence of electrolytes on the physical properties of the expansive soil. Then stabilized expansive soil with different percentage of electrolytes are tested for engineering properties, like permeability, compaction, unconfined compressive strength and shear strength properties to study the influence of electrolytes on expansive soil.

In this section the details of laboratory experimentation were presented. Analysis and discussion of test results will be presented in the next section.

III. RESULTS AND DISCUSSION

3.1. Effect of Additives on Atterberg's Limits

The variation of liquid limit values with different percentages of chemicals added to the expansive soil is presented in the Fig. 1. It is observed that the decrease in the liquid limit is significant upto 1% of chemical added to the expansive clay for all the chemicals, beyond 1% there is a nominal decrease. Maximum decrease in liquid limit for stabilized expansive clay is observed with the chemical FeCl₃, compared with other two chemicals, KCl and CaCl₂. Nominal increase in plastic limit of stabilized expansive clay is observed with increase the percentage of the chemical (Fig. 2).

Fig. 3 shows the variation of plasticity index with the addition of chemicals to expansive clay. The increase in the plastic limit and the decrease in the liquid limit cause a net reduction in the plasticity index. It is observed that, the reduction in plasticity indexes are 26%, 41% and 48% respectively for 1 % of KCl, CaCl₂ and FeCl₃ added to the expansive clay. The reduction in plasticity index with chemical treatment could be attributed to the depressed double layer thickness due to cation exchange by potassium, calcium and ferric ions.

The variation of shrinkage limit with the percentage of chemical added to the expansive soil is presented in the Fig. 4. With increase in percentage of chemical added to the expansive soil the shrinkage limit is increasing. With 1.5 % chemical addition, the shrinkage limit of stabilized

expansive clay is increased from 12% to 15.1%, 15.4% and 16% respectively for KCl, CaCl₂ and FeCl₃.

3.2. Effect of Additives on DFS

The variation of DFS of stabilized expansive clay with addition of different percentages of chemicals is shown in the Fig.5. It is observed that the DFS is decreasing with increasing percentage of chemical added to the expansive soil. Significant decrease in D.F.S. is recorded in stabilized expansive clay with addition of 1% of chemical. The reductions in the DFS of stabilized expansive clay with addition of 1% chemical are 40%, 43% and 47% for KCl, CaCl₂ and FeCl₃ respectively compared with the expansive clay. The reduction in DFS values could be supported by the fact that the double layer thickness is suppressed by cation exchange with potassium, calcium and ferric ions and with increased electrolyte concentration.

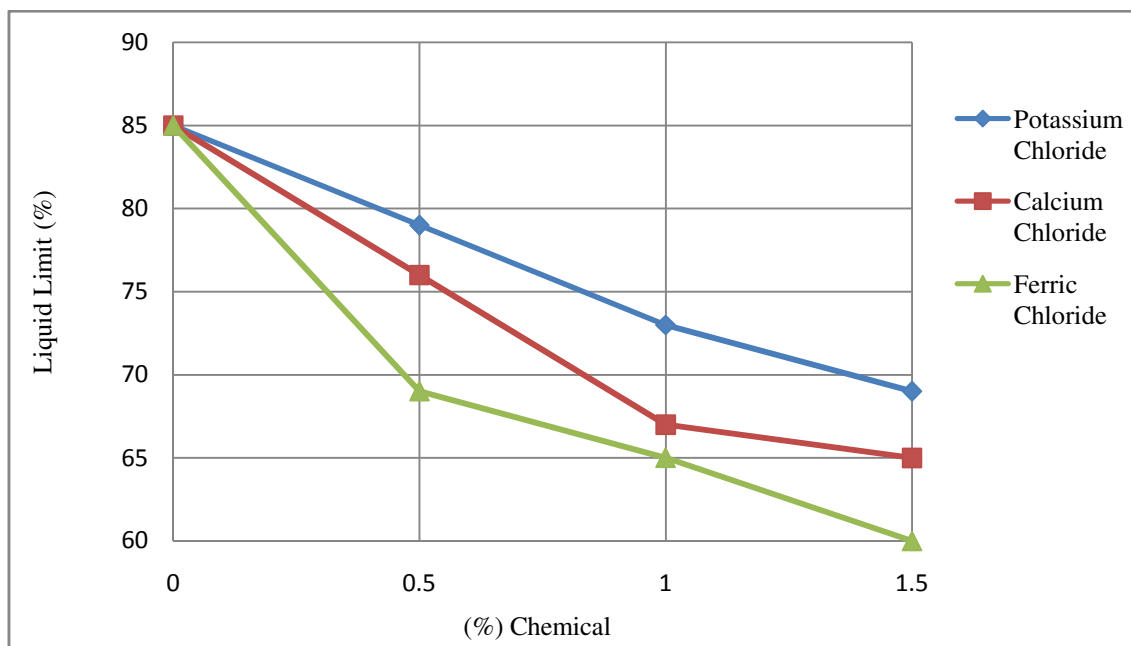


Fig.1: Variation of liquid limit with addition of percentage Chemical

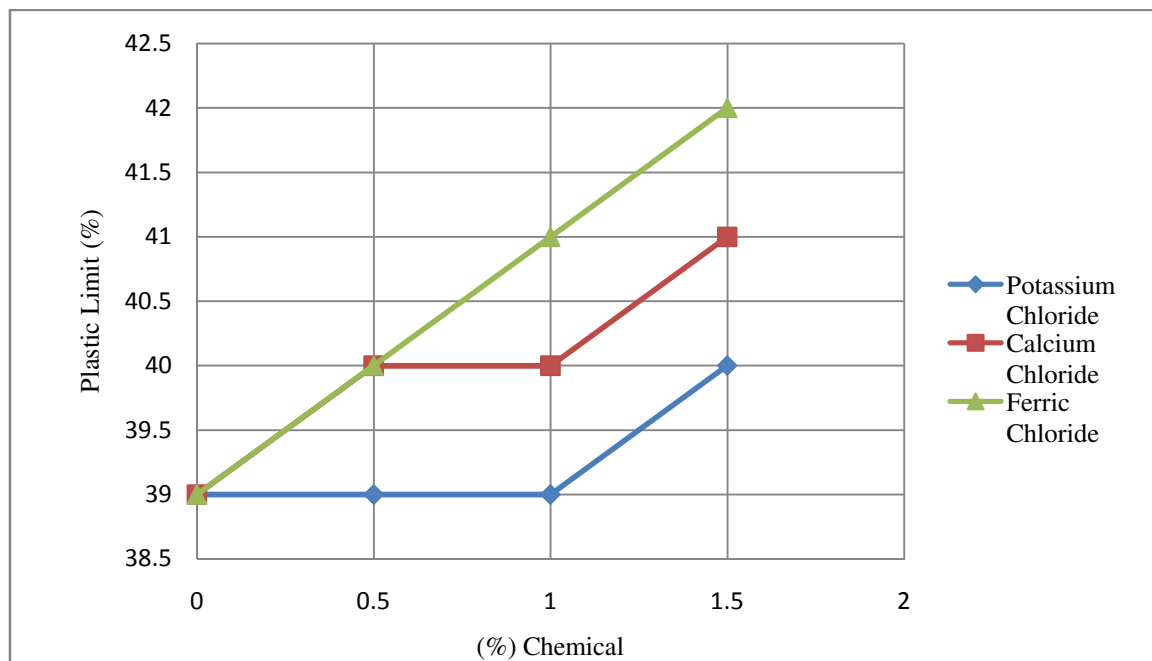


Fig.2: Variation of Plastic limit with addition of percentage Chemical

3.3. Effect of Additives on CBR

Fig. 6 shows the variation of CBR of stabilized expansive clay with addition of different percentages of chemicals. It can be seen that the CBR is increasing with increasing percentage of chemical added to the expansive soil. Significant increase in CBR is recorded in stabilized expansive clay with addition of chemical upto 1%, beyond this percentage the increase in CBR is marginal. The increase in CBR values of stabilized expansive clay with addition of 1% chemical are 80%, 99% and 116% for KCl, CaCl_2 and FeCl_3 respectively compared with the expansive clay. The increase in the strength with addition of chemicals may be attributed to the cation exchange of KCl, CaCl_2 & FeCl_3 between mineral layers and due to the formation of silicate gel. The reduction in improvement in CBR beyond 1% of chemicals KCl, CaCl_2 & FeCl_3 , may be due to the absorption of more moisture at higher chemical content.

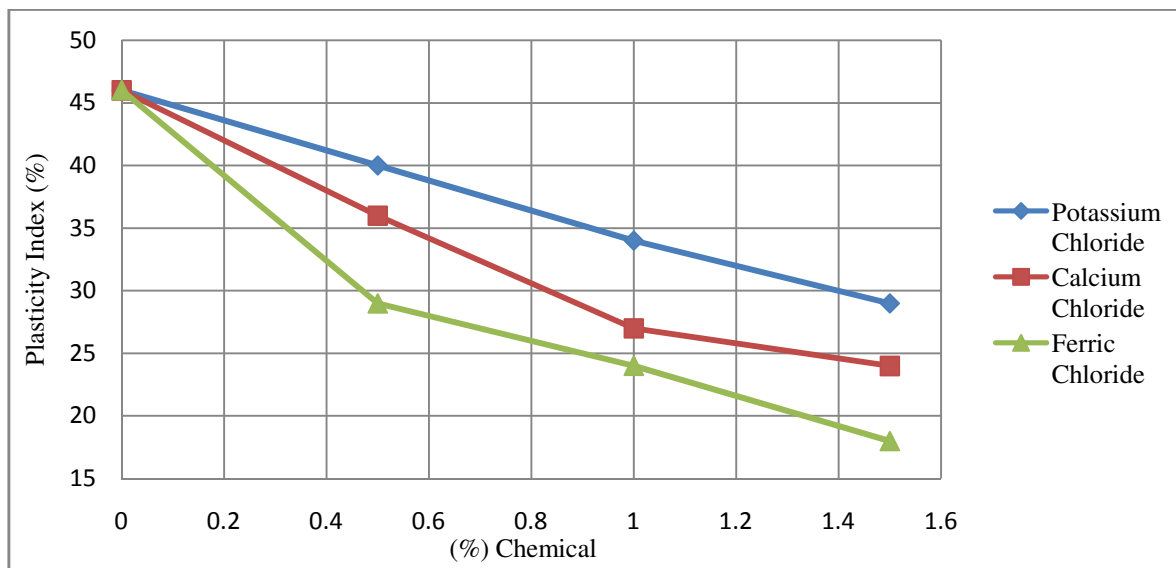


Fig.3: Variation of plasticity index with addition of percentage Chemical

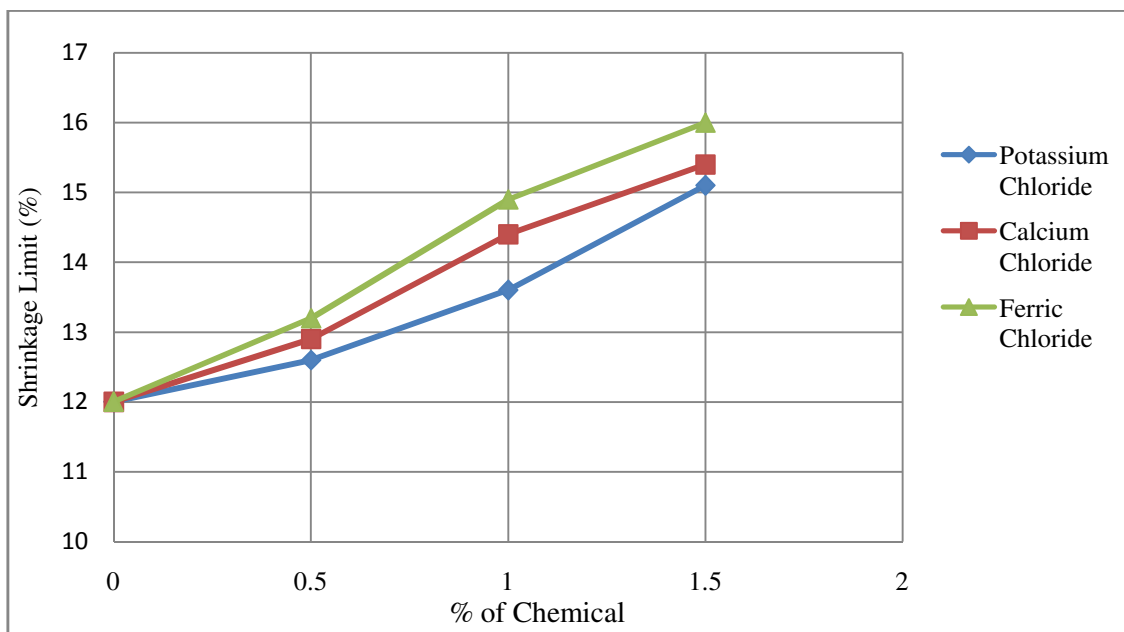


Fig.4: Variation of shrinkage limit with addition of percentage Chemical

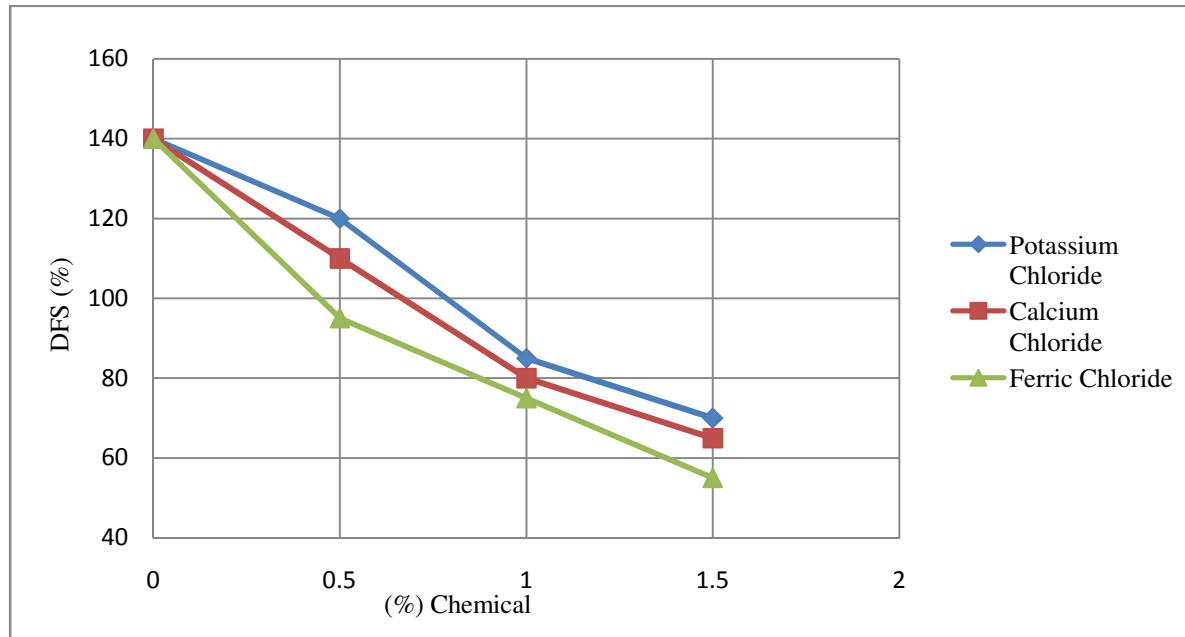


Fig. 5 Variation of DFS with addition of percentage Chemical

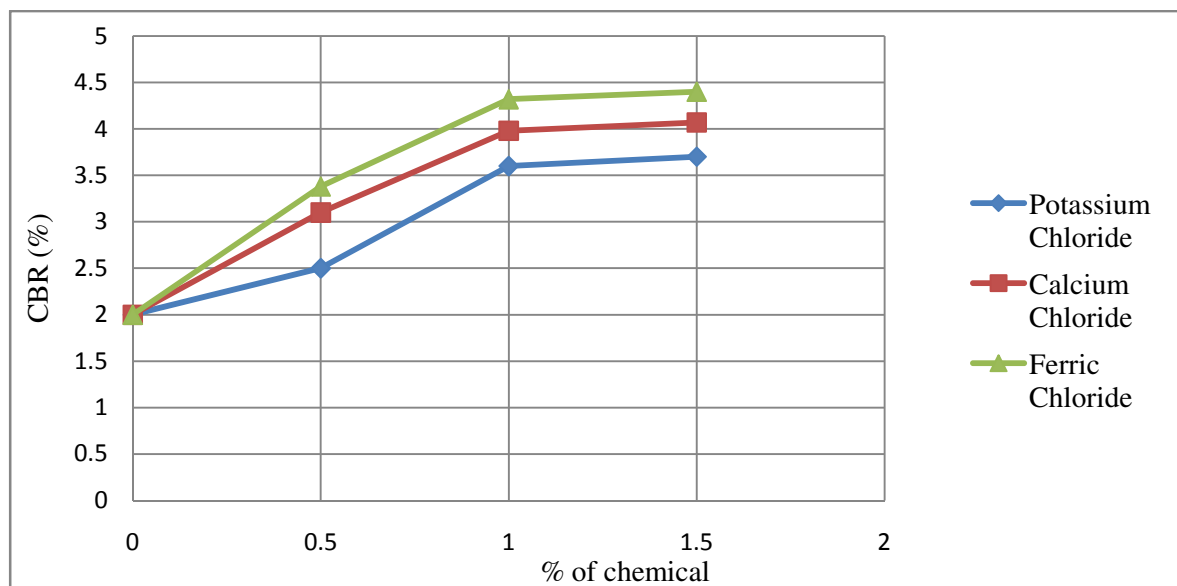


Fig.6: Variation of CBR of stabilized expansive bed with percentage of Chemical

3.4. Effect of Additives on Shear Strength Properties

The unconfined compressive strength of the remoulded samples prepared at MDD and optimum moisture content with addition of 0.5%, 1% and 1.5 % of chemicals, KCl, CaCl₂ & FeCl₃, to the expansive soil are presented in the table 2. The prepared samples are tested after 1day, 7 days and 14 days. As expected, the unconfined compressive strength is increasing with time may be due chemical reaction. It is observed that the unconfined compressive strength of the stabilized expansive soil is increasing with increase in percentage of chemical added to the soil. The unconfined compressive strength of stabilized expansive clay is increased by 133%, 171% & 230% when treated with 1% chemical, of KCl, CaCl₂ and FeCl₃ respectively. The increase in the strength with addition of chemicals may be attributed to the cation exchange of KCl, CaCl₂ & FeCl₃ between mineral layers and

due to the formation of silicate gel. The reduction in strength beyond 1% each of KCl, CaCl₂ & FeCl₃ may be due to the absorption of more moisture at higher KCl, CaCl₂ & FeCl₃.

The undrained shear strength parameters of the remoulded samples prepared at MDD and optimum moisture content with addition of 0.5%, 1% and 1.5 % of chemicals, KCl, CaCl₂ & FeCl₃, to the expansive soil are presented in the table 3. The prepared samples are tested after 1day, 7 days and 14 days. Significant change in undrained cohesion and marginal change in angle of internal friction is observed with addition of chemicals to the expansive clay. The increase in the shear strength parameters with addition of chemicals may be attributed to the cation exchange of chemicals. The shear strength parameters are increases upto 1 % chemical addition of above three chemicals, beyond this percentage there is a considerable decrease is observed may be due to the absorbtion of more moisture at higher chemical content.

Table: 2 Variation of Undrained compressive strength of stabilized expansive clay

Chemical added to the soil	Percentage of Chemical added to the soil	Unconfined Compressive Strength (KPa)		
		1 day	7 days	14days
Without chemical	--	92	--	--
KCl	0.5	130	175	188
	1.0	170	185	215
	1.5	125	160	180
CaCl ₂	0.5	135	200	215
	1.0	175	215	250
	1.5	128	184	207
FeCl ₃	0.5	140	245	256
	1.0	181	270	304
	1.5	132	223	248

Table: 3 Variation of Shear strength parameters with the addition of chemicals to the expansive clay

Chemical added to the soil	Percentage of Chemical added to the soil	Unconfined Compressive Strength (KPa)					
		1 day		7 days		14days	
		Cohesion, C _u (kg/cm ²)	Angle of internal friction, ϕ , (Deg.)	Cohesion, C _u (kg/cm ²)	Angle of internal friction, ϕ , (Deg.)	Cohesion, C _u (kg/cm ²)	Angle of internal friction, ϕ , (Deg.)
Without chemical	--	0.56	2 ⁰	--	--	--	--
KCl	0.5	0.61	7 ⁰	1.11	5 ⁰	1.28	7 ⁰
	1.0	0.72	5 ⁰	1.23	4 ⁰	1.32	4 ⁰
	1.5	0.65	6 ⁰	1.15	4 ⁰	1.26	4 ⁰
CaCl ₂	0.5	0.70	7 ⁰	1.21	5 ⁰	1.30	4 ⁰
	1.0	0.78	6 ⁰	1.32	5 ⁰	1.38	3 ⁰
	1.5	0.77	6 ⁰	1.27	4 ⁰	1.34	3 ⁰
FeCl ₃	0.5	0.89	6 ⁰	1.26	4 ⁰	1.33	3 ⁰
	1.0	0.96	4 ⁰	1.35	3 ⁰	1.46	3 ⁰
	1.5	0.93	3 ⁰	1.30	4 ⁰	1.38	3 ⁰

In this section the results of various tests carried out in the laboratory are discussed. Conclusions will be discussed in the next section.

IV. CONCLUSIONS

The following conclusions can be drawn from the laboratory study carried out in this investigation.

It is observed that the liquid limit values are decreased by 57 %, 63% and 70% respectively for 1% of KCl, CaCl₂ and FeCl₃ chemicals added to the expansive clay. Marginal increase in plastic limits is observed with addition of chemical to the expansive clay. Decrease in plasticity index is recorded with addition of chemical to the expansive soil. The shrinkage limit is increasing with 1.5 % chemical addition; it is observed that the shrinkage limit of stabilized expansive clay is increased from 12% to 15.1%, 15.4% and 16% respectively for KCl, CaCl₂ and FeCl₃.

The DFS values are decreased by 40%, 43% and 47% for 1% of KCl, CaCl₂ and FeCl₃ treatments respectively.

The CBR values are also increased by 80%, 103% and 116% respectively for 1% of KCl, CaCl₂ and FeCl₃ treatment.

It is observed that the Significant change in undrained cohesion and marginal change in angle of internal friction is observed with addition of chemicals to the expansive clay.

The UCS values are increased by 133%, 171% and 230% respectively for 1% of KCl, CaCl₂ and FeCl₃ treatments for a curing period of 14 day

REFERENCES

- [1] Babushanker N. (1986), "What Techniques other than under reamed piles have? Proven to be Effective in Minimizing Foundation Problems in Black Cotton Soils", IGC-86, New Delhi, Vol 1, pp.155-158.
- [2] Bansal, R.K., Pandey, P.K. and Singh, S.K (1996), "Improvement of a Typical Clay for Road Subgrades with Hydrated Lime", Proc. of National Conf. on Problematic Subsoil Conditions, Terzaghi-96, Kakinada, India, pp.193-197.
- [3] Bhattacharya, P. And Bhattacharya, A. (1989). "Stabilization of Bad banks of Railway Track by Lime Slurry Pressure Injection Technique", Proc. Of IGC-89, Visakhapatnam, Vol. 1, pp. 315-319.
- [4] Basma, A.A. and Tuncer, R.E. (1991). "Effect of Lime on Volume Change and Compressibility of Expansive Clays", TRR-1295, pp.52-61.
- [5] Bell, F.G. (1993). "Engg. Treatment of Soils", E&FN Spon Pub. Co.
- [6] Broms. B.B. and Boman, P. (1978). "Stabilization of soil with lime columns", Design hand book, second edition, royal institute of technology, Sweden.
- [7] Bredenberg, H. and Tekn, D.R. (1983). "Lime Columns for Ground improvement at New Cargo Terminal in Stockholm", Proc. Of the Eighty European Conf. on Soil Mechanics and Foundation Engg., Helsinki, pp. 881-884.
- [8] CRRI. (1978). "Handbook on Under-reamed and Bored Compaction Pile Foundation", Jain Printing Press, Roorkee, India.
- [9] Chen, F.H. (1988). "Foundations on Expansive Soils", Elsevier publications Co., Amsterdam.
- [10] Chummar, A.V. (1987). "Treatment of Expansive Soil Below Existing Structures with Sand – Lime Piles", Proc. Fo sixth Int. Conf. on expansive soils, New Delhi, pp. 451-452.
- [11] Desai, I.D. and Oza, B.N. (1977), "Influence of Anhydrous Calcium Chloride on the Shear Strength of Expansive soils, , Proc. of the First National Symposium on Expansion soils, HBTI-Kanpur, India, pp 4-1 to 4-5.
- [12] Deshpande, M.D. et al. (1990). "Performance Study of Road Section Constructed with Local Expansive Clay (Stabilized with lime) as Subbase material", Indian highways, pp. 35-41.
- [13] Frydman, S., Ravins, L and Ehrenreich, T. (1997), "Stabilization of Heavy Clay with Potassium Chloride", Journal of Geo technical Engg., Southeast Asian Society of Soil Engg., Vol 8, pp. 95-108.
- [14] Gichaga, F.J. (1991). "Deflections of Lateritic Gravel-Based and Stone Based Pavemetns of a Low-Volume Tea Road in Kenya", TRR-1291. TRB. Pp. 79-83.
- [15] Gokhale, K.V.G.K. (1977). "Mechanism of Soil Stabilization with Additives", Proc, of the first national symposium on expansive soils, HBTI, Kanpur, pp. 10-1 to 10-5.
- [16] Gokhale, Y.C. (1969). "Some Highway Engg. Problems in Black Cotton Soil Region", Proc. of the Symposium on characteristics of and construction techniques in black cotton soil pp, 209-212.
- [17] Gupta, A.K., Jain, S.S. and Bhatia, S.K. (1992). "A Study on Relationship between Rut Depth, Deflection and other Distress modes for flexible pavements", IRC Journal pp. 141-187.

- [18] Hausmann, M.R. (1990). "Engg . Principles of Ground Modification", Mc Graw Hill Book Co., New Delhi.
- [19] Ho, M.K (1968). "Swelling Characteristics of an Expansive Clay with Access to Common Electrolytes". Proc. of the Southeast Asian Regional Conf. on soil engg., Asian institute of Tech., Bangkok, pp. 159-167.
- [20] Holtz, W.G. (1969). "Volume Change in Expansive Clay Soils and Control by lime Treatment". Proc. of 2nd Int. Research and Engg. Conf on expansive clayey soils, Texas A & M Press, Texas, pp. 157-174.
- [21] Holtz, W.G. (1959): "Expansive Clays – Properties and Problems", First Annual Soil Mechanics Conf., Colorado School of Mines, Colorado, pp. 1-26.
- [22] Hopkins, T.C., Hunsucker, D.Q. and Beckam, T. (1994). "Selection of Design Strengths of Untreated Soil Sub grades and Sub grades treat with cement and hydrated lime". TRR-1440, TRB, pp. 37-44.
- [23] Holm, G., Brendenberg, H. and Broms, B.B. (1981): "Lime Columns as Foundation for Light Structures", Proc. of 10th ICSMFE, Stockholm, Vo. 3, pp. 687-694.
- [24] Humad, S. (1977). "Lime pile stabilization of Black cotton soil", Proc. of the 1st National Symposium on Expansive Soils, HBTI-Kanpur, India, pp. 4-1 to 4-8.
- [25] James, M.M. and Pandey, B.B (1990), "Performance of flexible pavements", TRR-1307, TRB, Washington, pp. 51-62.
- [26] Joshi, R.C., Natt, G.S. and Wright, P.J. (1981): "Soil improvement by lime-fly ash slurry injection", proc. of 10th Int. Conf. on IMFE, Vol. 3, Stockholm, PP. 707-712.
- [27] Katti, R.K., Kulkarni, K.R. and Radhakrishnan, N. (1966), "Research on Black Cotton Soils without and with Inorganic Additives", IRC, Road Research Bulletin, No. 10, pp. 1-97.
- [28] Koteswara Rao.D(2011), Laboratory investigations on GBFS- CH soil mixes for the utilization of foundation beds, CONCEPTS-2011, JNT university college of engineering, Kakinada.
- [29] Lakshmana Rao, C.B. et al. (1987) "Stabilization of Black cotton Soil with Inorganic Additives", Proc. of 6th Int. Conf. on expansive soils. New Delhi, India. Vol. 1, pp. 453-458.
- [30] Leonards, G.A. (1962). "Foundation Engineering", Mc-Graw Hill Book Co., New Delhi.
- [31] Little, N.D. (1996). "Assessment of In-situ Structural Properties of Lime Stabilized Clay Subgrades", TRR-1546, pp. 13-31.
- [32] Miller, G.A. and Zaman, M. (2000): "Field and laboratory evaluation of cement kiln dust as a soil stabilizer", TRR-1714, TRB, pp. 25-32.
- [33] Mitchell, J.K. and Radd, L. (1973). "Control of Volume Changes in Expansive Earth Materials", Proc. of workshop on expansive clays and shales in highway design and construction, Federal Highway Administration, Washington, D.C., pp. 200-217.
- [34] Mc Dowell, C. (1959). "Stabilization of Soils with Lime, Lime-flyash and other Lime reactive minerals", HRB, Bulletin No. 231.
- [35] Mohan Rai and jaisingh, M.P. (1985). "Advances in Building materials and construction", CBRI Publication, Roorkee, India.
- [36] O'Neil, M.W. and Poormoayed, N. (1980) "Methodology for foundation on expansive clays", journal of geo technical engg., proc. of ASCE, Vol. 106. No. GT 12.
- [37] Osinubi, K.J. (1988). "permeability of Lime Treated Lateritic Soil", Journal of transportation Eng., pp. 465-469.
- [38] Patel, A.N. and Quereshi, M.A., (1979). "A Methodology of Improving single lane roads in black cotton soil area", Indian Highways, pp. 5-11.
- [39] Petry, M.T. (1997). "Performance based testing of chemical stabilizers", TRR-1219, TRB, pp. 36-41.
- [40] Petry, T.M. and Armstrong, J.C. (1989), "Stabilization of Expansive Clay Soils", TRR-1219, TRB, pp. 103-112.
- [41] Pise, P.J. and Khanna, A.P. (1977): "Stabilization of Black Cotton Soil", Proc. of the first National Symposium on Expansive soils, HBTI, Kanpur, India, pp. 7-2 to 7-5.
- [42] Prasada Raju, G.V.R. (2001). "Evaluation of flexible pavement performance with reinforced and chemical stabilization of expansive soil sub grades". A Ph.D Thesis submitted to Kakatiya University, Warangal, (A.P.)
- [43] Ramannaiah, B.K., Sivananda, M and Satya Priya, (1972), "Stabilization of Black Cotton Soil with lime and Rice-Husk-Ash", 13th Annual General Body Meeting of Indian Geotechnical Society.
- [44] Ramana Murthy, V. (1998). "Study on swell pressure and method of controlling swell of expansive soil", Ph.D. Thesis, Kakatiya university, REC, Warangal.
- [45] Ramana Sastry, M.V.B. (1989). "Strengthening Subgrades of Roads in Deltaic Areas of Andhra Pradesh", Proc of IGC-89, Visakhapatnam, India Vol.1, pp 181-184.
- [47] Ramana Sastry, M.V.B., Srinivasulu Reddy, M and Gangaraju, Ch.P. (1986). "Comparative Study of Effect of Addition of Rice-Husk-Ash and Cinder-Ash to Soil-Lime Mixtures", Indian highways, Vol. 14, No. 8, pp 5-14.

- [48] Rao, S.M and Subba Rao,K.S. (1994), "Ground heave from Caustic soda solution spillage – A case study", Journal of soils and foundations, Japanese Society of soil Mech and foundation Engg,Vol. 34, No. 2, pp. 13-18.
- [49] Reddy, K.C., et al (1981). "Structural evaluation of sub-base courses", journal of Indian Roads congress, 42-2-14, paper No. 341, pp. 215-251.
- [50] Rolt, J. et al. (1987) . "Performance of a Full-scale pavement design experiment in jamica, TRR-1117, TRB, pp, 38-46.
- [51] Rogers, CDF and Glendenning. S. (1994). "Slope Stabilization using Lime Piles". TRR-1440, TRb, pp. 63-70.
- [52] Seeds, S.B.et al. (1999), "Development of performance related specifications for hot-Mix Asphalt Pavements through westrack", TR-1575, TRB,pp. 85-91.
- [53] Shepard, J.M. et al. (1991). "Full depth reclamation with calcium chloride", TRR-1295, TRB, pp.87-94.
- [54] Sivaguru, N., Reddy, K.C., Rajagopal, A.S., Veeraraghavan, A. and Justo, C.E.G. (1986). "Studies on New Flexible Pavements", IRC, Vol. 47-1, Paper No. 375 pp. 111-156.
- [55] Sivanna,G.S. et al. (1976). "Strength and consolidation characteristics of black cotton soil with chemical additives – CaCl₂ & KOH", report prepared by Karnataka Engg. Research station, Krsihnrajasagar,India.
- [56] Sivapullaiah, P.V. et al. (1994), "Role of electrolytes on the shear strength of clayey soil", Proc. of IGC-94, Warangal, pp. 199-202.
- [57] Slate, F.O. and Johnson, A.W. (1958), "Stabilization of soil with calcium chloride", HRB,Bibliography-24, pp. 1-90.
- [58] Snethen, D.R. et al. (1979), "An evaluation methodology for prediction and minimization of detrimental volume change of expansive soils in highway subgrades", research report, Vol. 1, prepared for federal highway administration, Washington.
- [59] Special Report -14 , IRC, (1995), "Ground Improvement by Lime stabilization". IRC, Highway research board, Washington.
- [60] Srinivas, M. (2008), Test track studies on chemically stabilized expansive soil subgrades, a Ph.D. thesis, JNT University college of engineering, Kakinada, June 2008.
- [61] Subba Rao, K.S. (1986). "What Techniques other than Under-Reamed pile have proven to be effective in minimizing foundation problem in black cotton soils", Proc of IGC-86, New Delhi, Vol. 1, pp. 149-154.
- [62] Subba Rao, K.S. (1999). "Swell – Shrink behavior of Expansive Slits-Geo technical Challenges", 22nd IGS Annual Lecture, IGC-99, Calcutta.
- [63] Thompson, M.R and Robnett, Q.L. (1976). "Pressure Injected Lime for Treatment of Swelling Soils", one of the 4 Reports prepared for the 54th Annual meeting of the TRB, TRR-568, pp 24-34.
- [64] Uppal, H.L. and Chadda,L.R. (1969). "Some Problems of Road Construction in Black Cotton Soils and the Remedial Measures". Proc of symposium on characteristics of and construction techniques in black cotton soil, the college of Military Engg., Puna, India, pp. 215-218.
- [65] Venkatanarayana, P. et al. (1989). "Ground Improvement by Sand – Lime Columns", Proc. Of IGC-89, Visakhapatnam, India, Vol. 1 , pp. 335-338.
- [66] Wright, P.J. (1973). "Lime Slurry Pressure Injection in Tames Expansive Clays", Civil Engg. ASCE.

Authors Biographies

P. Venkata Muthyalu is the post Graduate Student of Department of Civil Engg., University College of Engineering, JNTUK, Kakinada, India.



K. Ramu is working as Associate Professor in Department of Civil Engineering, JNTU College of Engineering, Kakinada, India. He has guided 15 M.Tech Projects & has 20 publications.



G.V.R. Prasada Raju is Professor of Civil Engineering and Director Academic Planning, JNTUK Kakinada, India. He is guiding 6 PhD scholars & 4 has been awarded the PhD. He has guided 60 M.Tech Projects & has 97 publications.



DESIGNING AN AUTOMATED SYSTEM FOR PLANT LEAF RECOGNITION

Jyotismita Chaki¹ and Ranjan Parekh²

School of Education Technology, Jadavpur University, Kolkata, India

ABSTRACT

This paper proposes an automated system for recognizing plant species based on leaf images. Plant leaf images corresponding to three plant types, are analyzed using three different shape modelling techniques, the first two based on the Moments-Invariant (M-I) model and the Centroid-Radii (C-R) model and the third based on a proposed technique of Binary-Superposition (B-S). For the M-I model the first four central normalized moments have been considered. For the C-R model an edge detector has been used to identify the boundary of the leaf shape and 36 radii at 10 degree angular separation have been used to build the shape vector. The proposed approach consists of comparing binary versions of the leaf images through superposition and using the sum of non-zero pixel values of the resultant as the feature vector. The data set for experimentations consists of 180 images divided into training and testing sets and comparison between them is done using Manhattan, Euclidean and intersection norms. Accuracies obtained using the proposed technique is seen to be an improvement over the M-I and C-R based techniques, and comparable to the best figures reported in extant literature.

KEYWORDS: Plant recognition, Moment Invariants, Centroid Radii, Binary Superposition, Computer Vision

I. INTRODUCTION

It is well known that plants play a crucial role in preserving earth's ecology and environment by maintaining a healthy atmosphere and providing sustenance and shelter to innumerable insect and animal species. Plants are also important for their medicinal properties, as alternative energy sources like bio-fuel and for meeting our various domestic requirements like timber, clothing, food and cosmetics. Building a plant database for quick and efficient classification and recognition of various flora diversities is an important step towards their conservation and preservation. This is more important as many types of plants are now at the brink of extinction. In recent times computer vision methodologies and pattern recognition techniques have been successfully applied towards automated systems of plant cataloguing. From this perspective the current paper proposes the design of a system which uses shape recognition techniques to recognize and catalogue plants based on the shape of their leaves, extracted from digital images. The organization of the paper is as follows : section 2 discusses an overview of related works, section 3 outlines the proposed approach with discussions on feature computation and classification schemes, section 4 provides details of the dataset and experimental results obtained, and section 5 brings up the overall conclusion and scopes for future research.

II. PREVIOUS WORKS

Many methodologies have been proposed to analyze plant leaves in an automated fashion. A large percentage of such works utilize shape recognition techniques to model and represent the contour shapes of leaves, however additionally, color and texture of leaves have also been taken into consideration to improve recognition accuracies. One of the earliest works [1] employs geometrical parameters like area, perimeter, maximum length, maximum width, elongation to differentiate between four types of rice grains, with accuracies around 95%. Use of statistical discriminant analysis

along with color based clustering and neural networks have been used in [2] for classification of a flowered plant and a cactus plant. In [3] the authors use the Curvature Scale Space (CSS) technique and k-NN classifiers to classify chrysanthemum leaves. Both color and geometrical features have been used in [4] to detect weeds in crop fields employing k-NN classifiers. In [5] the authors propose a hierarchical technique of representing leaf shapes by first their polygonal approximations and then introducing more and more local details in subsequent steps. Fuzzy logic decision making has been utilized in [6] to detect weeds in an agricultural field. In [7] the authors propose a two step approach of using a shape characterization function called centroid-contour distance curve and the object eccentricity for leaf image retrieval. The centroid-contour distance (CCD) curve and eccentricity along with an angle code histogram (ACH) have been used in [8] for plant recognition. The effectiveness of using fractal dimensions in describing leaf shapes has been explored in [9]. In contrast to contour-based methods, region-based shape recognition techniques have been used in [10] for leaf image classification. Wang et al. [11] describes a method based on fuzzy integral for leaf image retrieval. In [12] the authors used an improved CSS method and applied it to leaf classification with self intersection. Elliptic Fourier harmonic functions have been used to recognize leaf shapes in [13] along with principal component analysis for selecting the best Fourier coefficients. In [14] the authors propose a leaf image retrieval scheme based on leaf venation, represented using points selected by the curvature scale scope corner detection method on the venation image and categorized by calculating the density of feature points using non parametric estimation density. In [15] the authors propose a new classification method based on hypersphere classifier based on digital morphological feature. In [16] 12 leaf features are extracted and orthogonalized into 5 principal variables which consist of the input vector to a neural network (NN), trained by 1800 leaves to classify 32 kinds of plants. NNs have also been used in [17] to classify plants based on parameters like size, radius, perimeter, solidity and eccentricity of the leaf shape. In [18] shape representation is done using a new contour descriptor based on the curvature of the leaf contour. Wavelet and fractal based features have been used in [19] to model the uneven shapes of leaves. Texture features along with shape identifiers have been used in [20] to improve recognition accuracies. Other techniques like Zernike moments and Polar Fourier Transform have been proposed in [21] for modeling leaf structures. In [22] authors propose a thresholding method and H-maxima transformation based method to extract the leaf veins for vein pattern classification. In [23] authors propose an approach for combining global shape descriptors with local curvature-based features, for classifying leaf shapes of nearly 50 tree species. Finally in [24] a combination of all image features viz. color, texture and shape, have been used for leaf image retrieval.

III. SHAPE RECOGNITION TECHNIQUES

In this section we review two existing methods of shape recognition which have been used for plant classification, namely Moments-Invariant (M-I) and Centroid-Radii (C-R) and compare it with our proposed technique with respect to their recognition accuracies.

3.1. Moments Invariant (M-I) Approach: An Overview

M. K. Hu [25] proposes 7 moment features that can be used to describe shapes which are invariant to rotation, translation and scaling. For a digital image, the moment of a pixel $P(x, y)$ at location (x, y) is defined as the product of the pixel value with its coordinate distances i.e. $m = x.y.P(x, y)$. The moment of the entire image is the summation of the moments of all its pixels. More generally the moment of order (p, q) of an image $I(x, y)$ is given by

$$m_{pq} = \sum_x \sum_y [x^p y^q I(x, y)] \quad (1)$$

Based on the values of p and q the following are defined:

$$\begin{aligned}
m_{00} &= \sum_x \sum_y [x^0 y^0 I(x, y)] = \sum_x \sum_y [I(x, y)] \\
m_{10} &= \sum_x \sum_y [x^1 y^0 I(x, y)] = \sum_x \sum_y [x I(x, y)] \\
m_{01} &= \sum_x \sum_y [x^0 y^1 I(x, y)] = \sum_x \sum_y [y I(x, y)] \\
m_{11} &= \sum_x \sum_y [x^1 y^1 I(x, y)] = \sum_x \sum_y [xy I(x, y)] \\
m_{20} &= \sum_x \sum_y [x^2 y^0 I(x, y)] = \sum_x \sum_y [x^2 I(x, y)] \\
m_{02} &= \sum_x \sum_y [x^0 y^2 I(x, y)] = \sum_x \sum_y [y^2 I(x, y)] \\
m_{21} &= \sum_x \sum_y [x^2 y^1 I(x, y)] = \sum_x \sum_y [x^2 y I(x, y)] \\
m_{12} &= \sum_x \sum_y [x^1 y^2 I(x, y)] = \sum_x \sum_y [xy^2 I(x, y)] \\
m_{30} &= \sum_x \sum_y [x^3 y^0 I(x, y)] = \sum_x \sum_y [x^3 I(x, y)] \\
m_{03} &= \sum_x \sum_y [x^0 y^3 I(x, y)] = \sum_x \sum_y [y^3 I(x, y)]
\end{aligned} \tag{2}$$

The first four Hu invariant moments which are invariant to rotation are defined as follows:

$$\begin{aligned}
\phi_1 &= m_{20} + m_{02} \\
\phi_2 &= (m_{20} - m_{02})^2 + (2m_{11})^2 \\
\phi_3 &= (m_{30} - 3m_{12})^2 + (3m_{21} - m_{03})^2 \\
\phi_4 &= (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2
\end{aligned} \tag{3}$$

To make the moments invariant to translation the image is shifted such that its centroid coincides with the origin of the coordinate system. The centroid of the image in terms of the moments is given by:

$$\begin{aligned}
x_c &= \frac{m_{10}}{m_{00}} \\
y_c &= \frac{m_{01}}{m_{00}}
\end{aligned} \tag{4}$$

Then the central moments are defined as follows:

$$\mu_{pq} = \sum_x \sum_y [(x - x_c)^p (y - y_c)^q I(x, y)] \tag{5}$$

To compute Hu moments using central moments the ϕ terms in equation (3) need to be replaced by μ terms. It can be verified that $\mu_{00} = m_{00}$, $\mu_{10} = 0 = \mu_{01}$.

To make the moments invariant to scaling the moments are normalized by dividing by a power of μ_{00} . The *normalized central moments* are defined as below :

$$M_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\omega}, \text{ where } \omega = 1 + \frac{p+q}{2} \tag{6}$$

3.2. Centroid-Radii (C-R) Approach: An Overview

In [26] K. L. Tan et al. proposes the centroid-radii model for estimating shapes of objects in images. A shape is represented by an area of black on a white background. Each pixel is represented by its color (black or white) and its x - y co-ordinates on the canvas. The boundary of a shape consists of a series of boundary points. A *boundary point* is a black pixel with at least one white pixel as its

neighbor. Let (x_i, y_i) , $i = 1, \dots, n$ represent the shape having n boundary points. The *centroid* is located at the position $C(X_C, Y_C)$ which are respectively, the average of the x and y co-ordinates for all black pixels :

$$X_C = \frac{\sum_{i=1}^n x_i}{n}$$

$$Y_C = \frac{\sum_{i=1}^n y_i}{n}$$
(7)

A *radius* is a straight line joining the centroid to a boundary point. In the centroid-radii model, lengths of a shape's radii from its centroid to the boundary are captured at regular intervals as the shape's descriptor using the Euclidean distance. More formally, let θ be the measure of the angle (in degrees) between radii (Figure 1). Then, the number of angular intervals is given by $k = 360/\theta$. The length L_i of the i -th radius formed by joining the centroid $C(X_C, Y_C)$ to the i -th sample point $s_i(x_i, y_i)$ is given by:

$$L_i = \sqrt{(X_C - x_i)^2 + (Y_C - y_i)^2}$$
(8)

All radii lengths are normalized by dividing with the longest radius length from the set of radii lengths extracted. Let the individual radii lengths be $L_1, L_2, L_3, \dots, L_k$ where k is total number of radii drawn at an angular separation of θ . If the maximum radius length is L_{\max} then the normalized radii lengths are given by:

$$\ell_i = \frac{L_i}{L_{\max}}, \quad i = 1, \dots, k$$
(9)

Furthermore, without loss of generality, suppose that the intervals are taken clockwise starting from the x -axis direction (0°). Then, the shape descriptor can be represented as a vector consisting of an ordered sequence of normalized radii lengths:

$$S = \{\ell_0, \ell_\theta, \ell_{2\theta}, \dots, \ell_{(k-1)\theta}\}$$
(10)

With sufficient number of radii, dissimilar shapes can be differentiated from each other.

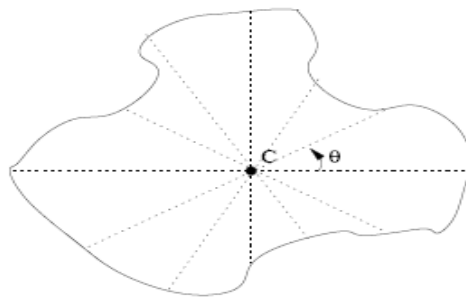


Figure 1. Centroid-radii approach

3.3. Proposed Approach: Binary Superposition (B-S)

The proposed approach is conceptually simpler than either of the above two techniques but provides comparatively better recognition accuracies. The leaf images are converted to binary images by thresholding with an appropriate value. Two binary shape images S_1 and S_2 are superimposed on each other and a resultant R is computed using logical AND operation between them.

$$R = S_1 \cap S_2$$
(11)

For the binary resultant image, all the non-zero pixel values are summed up. This sum is used as the feature value for discrimination. A large value of the sum would indicate high similarity between the images while a low sum value indicates low similarity. A test image is compared to all the training samples of each class and the mean resultant for each class is computed. The test image is classified to the class for which the mean resultant is maximum. Figure 2 illustrates that when two images of different classes are superimposed then the resultant image contains less non-zero pixels than for images belonging to the same class.

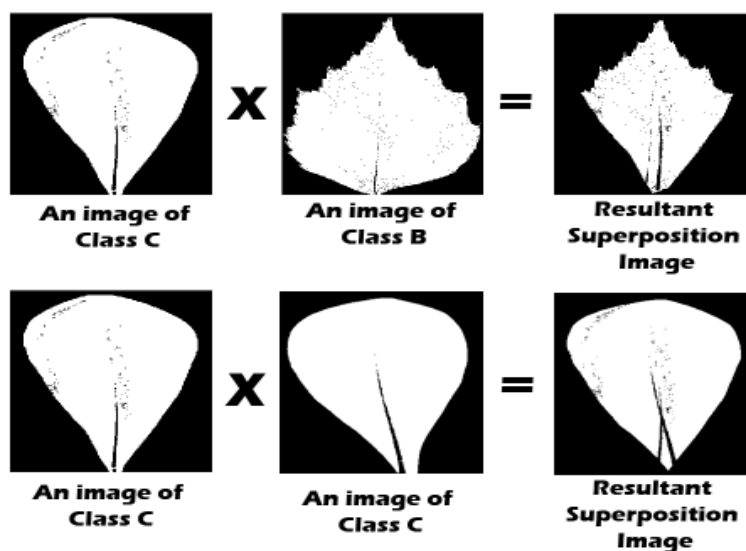


Figure 2. Resultant images after binary superposition

IV. EXPERIMENTATIONS AND RESULTS

4.1. Dataset

Experimentations are performed by using 180 leaf images from the Plantscan database [27]. The dataset is divided into 3 classes: A (*Arbutus unedo*), B (*Betula pendula* Roth), C (*Pittosporum tobira*) each consisting of 60 images. Each image is 350 by 350 pixels in dimensions and in JPEG format. A total of 120 images are used as the Training set (T) and the remaining 120 images as the Testing set (S). The legends used in this work are as follows : AT, BT, CT denotes the training set while AS, BS, CS denotes the testing set, corresponding to the three classes. Sample images of each class are shown below in Figure. 3.

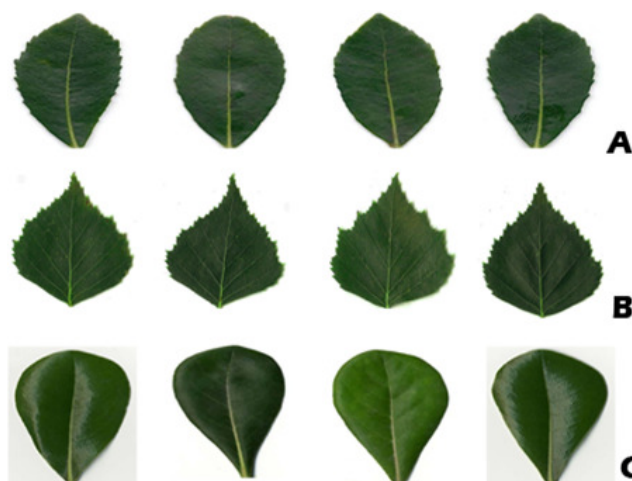


Figure 3. Samples of leaf images belonging to the three classes

4.2. M-I based computations

The first four moments M_1 to M_4 of each image of the training and testing sets are computed as per equation (6). Feature values are first considered individually and corresponding results are tabulated below. Comparisons between training and testing sets are done using Manhattan distances (L_1). Results are summarized below in Table 1. The last column depicts the overall percentage accuracy value.

Table 1 : Percentage accuracy using M-I approach

F	A	B	C	O
M_1	96	100	47	81
M_2	60	37	90	62
M_3	53	100	37	63
M_4	77	40	70	62

Best results of 81% are obtained using M_1 . Corresponding plots depicting the variation of the M_1 feature values for the three classes over the training and testing datasets are shown below in Figure 4.

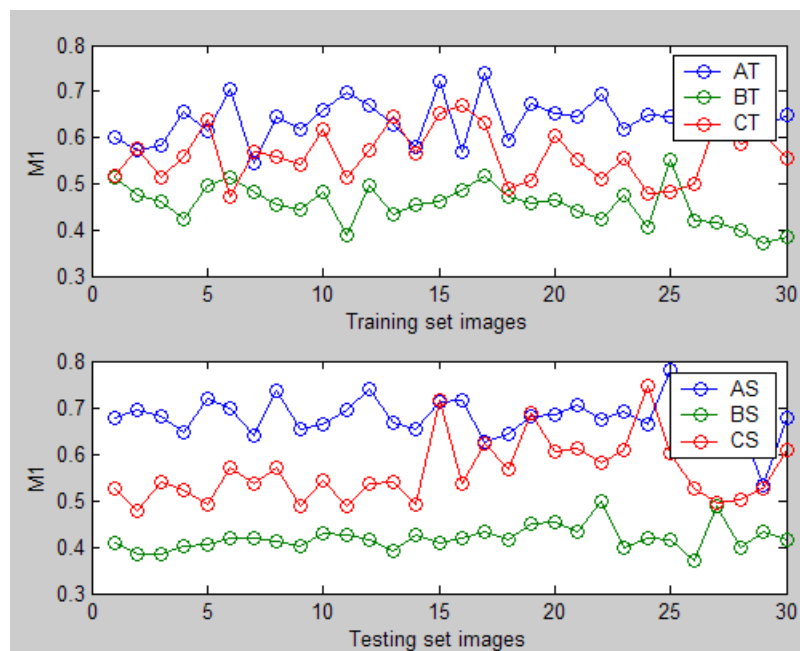


Figure 4. Variation of M_1 for Training and Testing set images

4.3. C-R based computations

Each image is converted to binary form and the Canny edge detector is used to identify its contour. Its centroid is computed from the average of its edge pixels. Corresponding to each edge pixel the angle it subtends at the centroid is calculated and stored in an array along with its x - and y - coordinate values. From the array 36 coordinate values of edge pixels which join the centroid at 10 degree intervals from 0 to 359 degrees are identified. The radii length of joining these 36 points with the centroid is calculated using the Euclidean distance and the radii lengths are normalized to the range [0, 1]. For each leaf image 36 such normalized lengths are stored in an ordered sequence as per equation (10). Figure 5 shows a visual representation of a leaf image, the edge detected version, the location of the centroid and edge pixels, and the normalized radii vector.

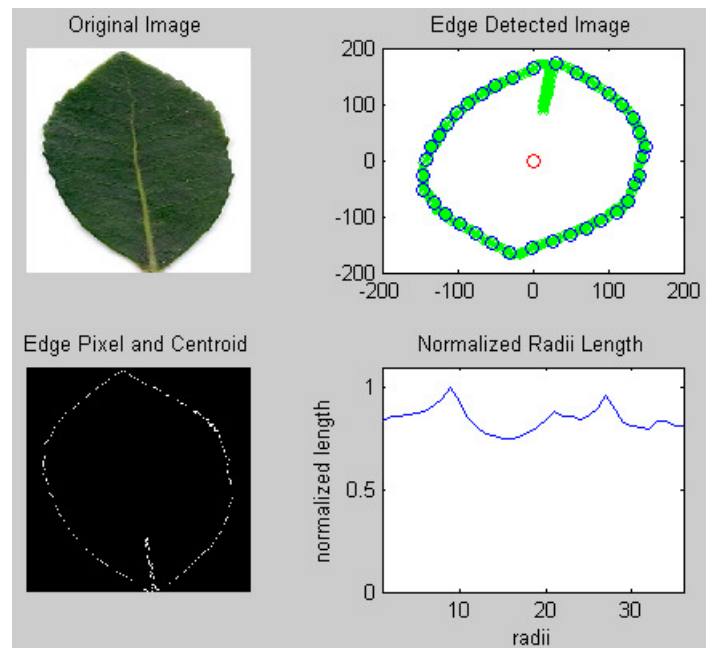


Figure 5. Interface for Centroid-Radii computation

The average of the 36 radii lengths for each image of each class both for the training and testing sets, is plotted in Figure 6, to depict the overall feature range and variation for each class.

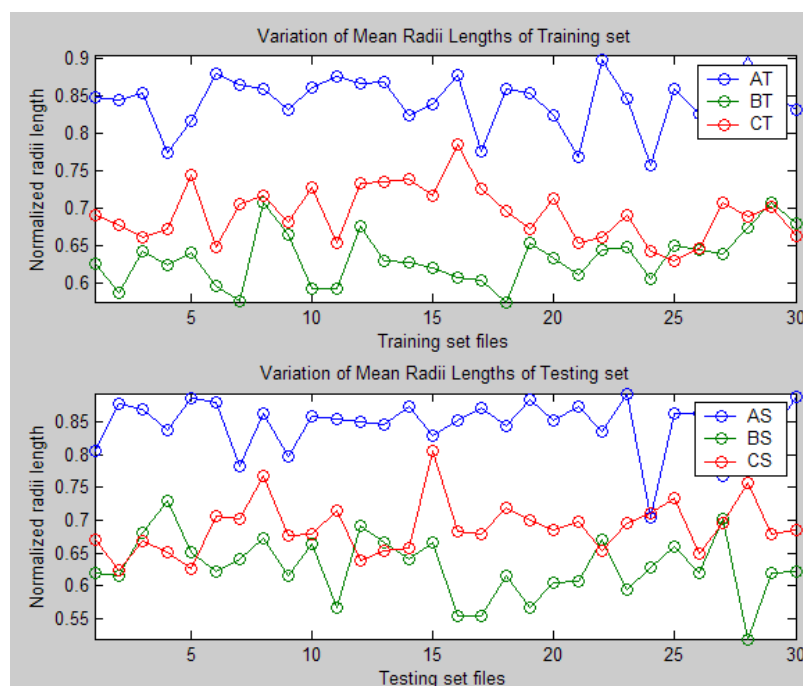


Figure 6. Variation of mean radii length for Training and Testing set images

Classes are discriminated using Euclidean distance (L_2) metric between the 36-element C-R vectors of training and testing samples. Results are summarized below in Table 2.

Table 2 : Percentage accuracy using C-R approach

F	A	B	C	O
C-R	97	100	97	98

4.4. B-S based computations

Each image is converted to binary form by using a 50% thresholding. A binarized test image is multiplied with each of the binarized training samples for each class as per equation (11) and the sum of the '1s' in the resultant image is used as the feature vector for discrimination. Figure 7 shows the variation of feature values for the three classes. The upper plot is obtained by binary superposition of Class-A testing images with all training samples, the middle plot is obtained by binary superposition of Class-B testing images with all training samples and the lower plot is obtained by binary superposition of Class-C testing images with all training samples.

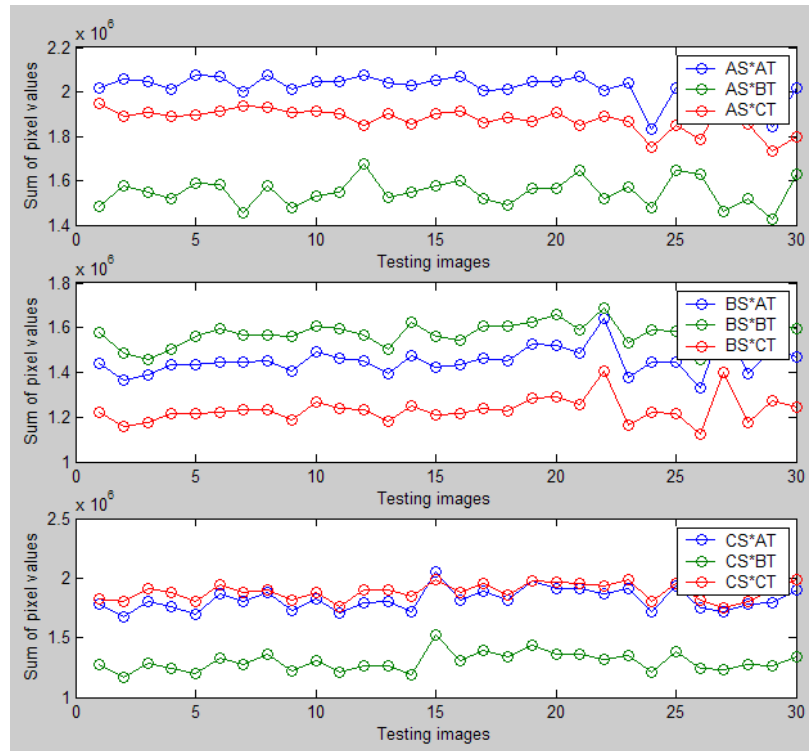


Figure 7. Variation of feature value for Testing set images using B-S approach

Classes are discriminated by determining the maximum values of the resultant B-S matrices computed by superimposing the training and testing samples. Results are summarized below in Table 3.

Table 3 : Percentage accuracy using B-S approach

F	A	B	C	O
B-S	100	100	97	99

V. ANALYSIS

Automated discrimination between three leaf shapes was done using a variety of approaches to find the optimum results. The study reveals that for M-I approach provide best results for the M1 feature. Accuracies based on C-R method using a 36-element radii vector provide results better than individual M-I features. The proposed approach of binary superposition improved upon the results provided by the M-I approach. Best results obtained using different methods are summarized below in Table 4.

Table 4 : Summary of accuracy results

F	A	B	C	O
M-I	96	100	47	81
C-R	97	100	97	98
B-S	100	100	97	99

To put the above results in perspective with the state of the art, the best results reported in [8] is a recall rate of 60% for discrimination of chrysanthemum leaves from a database of 1400 color images. Accuracy for classification for 10 leaf categories over 600 images is reported to be 82.33% in [10]. Overall classification accuracy reported in [11] for 4 categories of leaf images obtained during three weeks of germination, is around 90%. Accuracy reported in [13] for classification of 32 leaf types from a collection of 1800 images is around 90%. An overall classification of 80% is reported in [14] for identifying two types of leaf shapes from images taken using different frequency bands of the spectrum. Best accuracies reported in [17] are around 93% using Polar Fourier Transforms. Results reported in [20] are in the region of 80% for classifying 50 species. Accuracies of around 97% have been reported in [21] for a database of 500 images. It therefore can be said that the accuracies reported in the current paper are comparable to the best results reported in extant literature. It may however be noted that in many of the above cases color and geometrical parameters have also been combined with shape based features to improve results, while the current work is based solely on shape characteristics.

VI. CONCLUSIONS AND FUTURE SCOPES

This paper proposes an automated system for plant identification using shape features of their leaves. Two shape modelling approaches are discussed: one technique based on M-I model and the other on C-R model, and these are compared with a proposed approach based on binary superposition. The feature plots and recognition accuracies for each of the approaches are studied and reported. Such automated classification systems can prove extremely useful for quick and efficient classification of plant species. The accuracy of the current proposed approach is comparable to those reported in contemporary works. A salient feature of the proposed approach is the low-complexity data modelling scheme used whereby the computations only involve binary values.

Future work would involve research along two directions: (1) combining other shape based techniques like Hough transform and Fourier descriptors, and (2) combining color and texture features along with shape features for improving recognition accuracies.

REFERENCES

- [1] N. Sakai, S. Yonekawa, & A. Matsuzaki (1996), "Two-dimensional image analysis of the shape of rice and its applications to separating varieties", *Journal of Food Engineering*, vol 27, pp. 397-407.
- [2] A. J. M. Timmermans, & A. A. Hulzebosch (1996), "Computer vision system for on-line sorting of pot plants using an artificial neural network classifier", *Computers and Electronics in Agriculture*, vol. 15, pp. 41-55.
- [3] S. Abbasi, F. Mokhtarian, & J. Kittler (1997), "Reliable classification of chrysanthemum leaves through curvature scale space", *Lecture Notes in Computer Science*, vol. 1252, pp. 284-295.
- [4] A. J. Perez, F. Lopez, J. V. Benlloch, & S. Christensen (2000), "Color and shape analysis techniques for weed detection in cereal fields", *Computers and Electronics in Agriculture*, vol. 25, pp. 197-212.
- [5] C. Im, H. Nishida, & T. L. Kunii (1998), "A hierarchical method of recognizing plant species by leaf shapes", *IAPR Workshop on Machine Vision Applications*, pp. 158-161.
- [6] C-C Yang, S. O. Prasher, J-A Landry, J. Perret, and H. S. Ramaswamy (2000), "Recognition of weeds with image processing & their use with fuzzy logic for precision farming", *Canadian Agricultural Engineering*, vol. 42, no. 4, pp. 195-200.
- [7] Z. Wang, Z. Chi, D. Feng, & Q. Wang (2000), "Leaf image retrieval with shape feature", *International Conference on Advances in Visual Information Systems (ACVIS)*, pp. 477-487.
- [8] Z. Wang, Z. Chi, & D. Feng (2003), "Shape based leaf image retrieval", *IEEE Proceedings on Vision, Image and Signal Processing (VISIP)*, vol. 150, no.1, pp. 34-43.
- [9] J. J. Camarero, S. Siso, & E.G-Pelegrin (2003), "Fractal dimension does not adequately describe the complexity of leaf margin in seedlings of Quercus species", *Anales del Jardín Botánico de Madrid*, vol. 60, no. 1, pp. 63-71.
- [10] C-L Lee, & S-Y Chen (2003), "Classification of leaf images", *16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP)*, pp. 355-362.
- [11] Z. Wang, Z. Chi, & D. Feng (2002), "Fuzzy integral for leaf image retrieval", *IEEE Int. Conf. on Fuzzy Systems*, pp. 372-377.
- [12] F. Mokhtarian, & S. Abbasi (2004), "Matching shapes with self-intersection: application to leaf classification", *IEEE Trans. on Image Processing*, vol. 13, pp. 653-661.

- [13] J. C. Neto, G. E. Meyer, D. D. Jones, & A. K. Samal (2006), "Plant species identification using elliptic Fourier leaf shape analysis", *Computers and Electronics in Agriculture*, vol. 50, pp. 121-134.
- [14] J-K Park, E-J Hwang, & Y. Nam (2006), "A vention – based leaf image classification scheme", *Alliance of Information and Referral Systems*, pp. 416-428.
- [15] J. Du, X. Wang, & G. Zhang (2007), "Leaf shape based plant species recognition", *Applied Mathematics and Computation*, vol. 185, pp. 883-893.
- [16] S. G. Wu, F. S. Bao, E. Y. Xu, Y-X Wang, Y-F Chang, & Q-L Xiang (2007), "A leaf recognition algorithm for plant classification using probabilistic neural network", *The Computing Research Repository (CoRR)*, vol.1, pp. 11-16.
- [17] J. Pan, & Y. He (2008), "Recognition of plants by leaves digital image and neural network", *International Conference on Computer Science and Software Engineering*, vol 4, pp. 906 – 910.
- [18] C. Caballero, & M. C. Aranda (2010), "Plant species identification using leaf image retrieval", *ACM Int. Conf. on Image and Video Retrieval (CIVR)*, pp. 327-334.
- [19] Q-P Wang, J-X Du, & C-M Zhai (2010), "Recognition of leaf image based on ring projection wavelet fractal feature", *International Journal of Innovative Computing, Information and Control*, pp. 240-246.
- [20] T. Beghin, J. S. Cope, P. Remagnino, & S. Barman (2010), "Shape and texture based plant leaf classification", *International Conference on Advanced Concepts for Intelligent Vision Systems (ACVIS)*, pp. 345-353.
- [21] A. Kadir, L.E. Nugroho, A. Susanto, & P.I. Santosa (2011), "A comparative experiment of several shape methods in recognizing plants", *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 3, no. 3, pp. 256-263
- [22] N. Valliammal, & S. N. Geethalakshmi (2011), "Hybrid image segmentation algorithm for leaf recognition and characterization", *International Conference on Process Automation, Control and Computing (PACC)*, pp. 1-6.
- [23] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, & D. Coquin (2011), "Guiding active contours for tree leaf segmentation and identification", *Cross-Language Evaluation Forum (CLEF)*, Amsterdam, Netherlands.
- [24] B. S. Bama, S. M. Valli, S. Raju, & V. A. Kumar (2011), "Conten based leaf image retrieval using shape, color and texture features", *Indian Journal of Computer Science and Engineering*, vol. 2, no. 2, pp. 202-211.
- [25] M-K Hu (1962), "Visual pattern recognition by moment invariants", *IRE Transactions on Information Theory*, pp. 179-187.
- [26] K-L Tan, B. C. Ooi, & L. F. Thiang (2003), "Retrieving similar shapes effectively and efficiently", *Multimedia Tools and Applications*, vol. 19, pp. 111-134
- [27] Index of /Pl@ntNet/plantscan_v2 (http://imediapt.inria.fr:50012/PL@ntNet/plantscan_v2)

Authors

Jyotismita Chaki is a Masters (M.Tech.) research scholar at the School of Education Technology, Jadavpur University, Kolkata, India. Her research interests include image processing and pattern recognition.



Ranjan Parekh is a faculty at the School of Education Technology, Jadavpur University, Kolkata, India. He is involved with teaching subjects related to multimedia technologies at the post-graduate level. His research interests include multimedia databases, pattern recognition, medical imaging and computer vision. He is the author of the book "Principles of Multimedia" published by McGraw-Hill, 2006.



FUZZY CONTROL OF SQUIRREL CAGE INDUCTION MACHINE WIND GENERATION SYSTEM

B. Ravichandra Rao and R. Amala Lolly
Department of EEE Engineering, GNITS, Hyderabad, India

ABSTRACT

Artificial intelligence techniques, such as fuzzy logic, neural network and genetic algorithm are recently showing a lot of promise in the application of power electronic systems. This Paper describes the control strategy development, design and of a fuzzy logic based variable speed wind generation system. In this work cage type induction generator and double-sided PWM converters are used. The fuzzy logic based control of the system helps to optimize the efficiency and enhance the performance. The generation system uses three fuzzy logic controllers. The first fuzzy controller tracks the generator speed with the wind velocity to extract maximum power. The second fuzzy logic controller programs machine flux for light load efficiency improvement. The third fuzzy logic controller provides robust speed control against wind vortex and turbine oscillatory torque. The complete control system has been developed, analyzed, and simulated in Matlab.

KEYWORDS: Induction Generator, Fuzzy Logic Controller and Wind Generation system.

I. INTRODUCTION

GRID-connected wind electricity generation is showing the highest rate of growth of any form of electricity generation, achieving global annual growth rates in the order of 20 - 25%. Wind power is increasingly being viewed as a mainstream electricity supply technology. Its attraction as an electricity supply source has fostered ambitious targets for wind power in many countries around the world.

Wind power penetration levels have increased in electricity supply systems in a few countries in recent years; so have concerns about how to incorporate this significant amount of intermittent, uncontrolled and non-dispatchable generation without disrupting the finely-tuned balance that network systems demand.

Grid integration issues are a challenge to the expansion of wind power in some countries. Measures such as aggregation of wind turbines, load and wind forecasting and simulation studies are expected to facilitate larger grid penetration of wind power. In this project simulation studies on grid connected wind electric generators (WEG) employing Squirrel Cage Induction Generator (SCIG)[2].

Fuzzy Logic is a powerful and versatile tool for representing imprecise, ambiguous and vague information. It also helps to model difficult, even intractable problems. The system uses three fuzzy controllers, Fuzzy Programming of Generator Speed, Fuzzy Programming of Generator Flux, Fuzzy Control of Generator Speed Loop

II. WIND - GENERATION SYSTEM DESCRIPTION

2.1 Converter System

The AC/DC/AC converter is divided into two components: the rotor-side converter (C_{rotor}) and the grid-side converter (C_{grid}). C_{rotor} and C_{grid} are Voltage-Sourced Converters that use forced-commutated

power electronic devices (IGBTs) to synthesize an AC voltage from a DC voltage source. A capacitor connected on the DC side acts as the DC voltage source. A coupling inductor L is used to connect C_{grid} to the grid. The three-phase rotor winding is connected to C_{rotor} by slip rings and brushes and the three-phase stator winding is directly connected to the grid. The power captured by the wind turbine is converted into electrical power by the induction generator and it is transmitted to the grid by the stator and the rotor windings. The control system generates the pitch angle command and the voltage command signals V_r and V_{gc} for C_{rotor} and C_{grid} respectively in order to control the power of the wind turbine, the DC bus voltage and the reactive power or the voltage at the grid terminals.

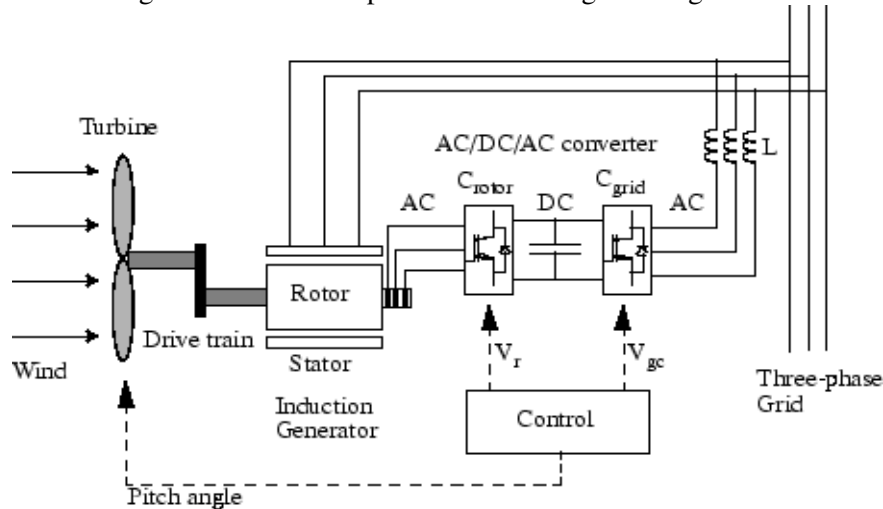


Fig1: The Wind Turbine Doubly-Fed Induction Generator

Lastly the generation system feeds power to a utility grid. Some of its salient features are as follows:

- Line side power factor is unity with no harmonic current injection.
- The cage type induction machine is extremely rugged, reliable, economical, and universally popular.
- Machine current is sinusoidal and no harmonic copper loss.
- Rectifier can generate programmable excitation for the machine.
- Continuous power generation from zero to highest turbine speed is possible.
- Power can flow in either direction permitting the generator to run as a motor for start-up.
- Autonomous operation is possible either with the help of start up capacitor or dc link battery.
- Extremely fast transient is also possible.

The mechanical power and the stator electric power output are computed as follows:

$$P_m = T_m \omega_r \text{---(1)}$$

$$P_s = T_{em} \omega_s \text{---(2)}$$

For a lossless generator the mechanical equation is:

$$J \frac{d\omega_r}{dt} = T \text{---(3)}$$

In steady-state at fixed speed for a lossless generator $T_m = T_{em}$ and $P_m = P_s + P_r$.

It follows that:

$$P_r = P_m - P_s = T_m \omega_r - T_{em} \omega_s = -T_m \frac{\omega_s - \omega_r}{\omega_s} \omega_s$$

$$P_r = -s T_m \omega_s$$

$$P_r = -s P_s \text{---(4)}$$

where s is defined as the slip of the generator: $s = (\omega_s - \omega_r) / \omega_s \text{---(5)}$

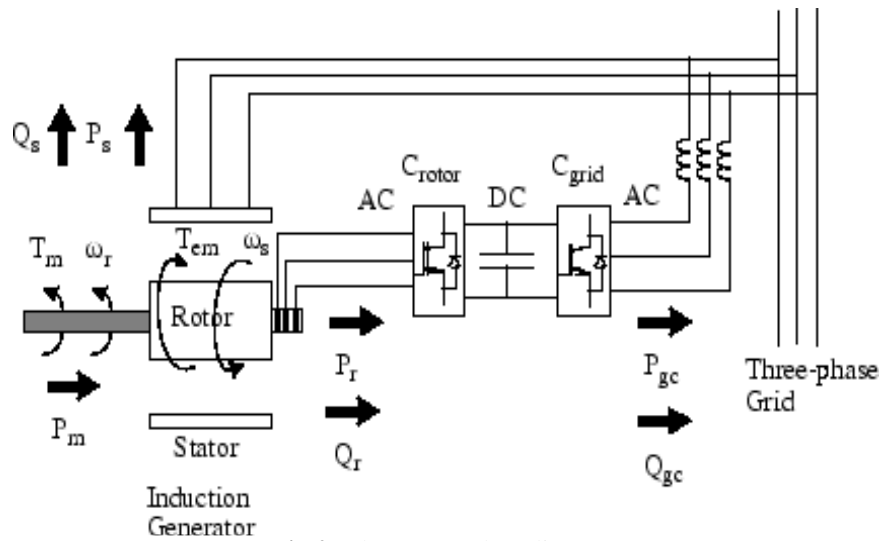


Fig 2: The Power Flow diagram

Generally the absolute value of slip is much lower than 1 and, consequently, P_r is only a fraction of P_s . Since T_m is positive for power generation and since ω_s is positive and constant for a constant frequency grid voltage, the sign of P_r is a function of the slip sign. P_r is positive for negative slip (speed greater than synchronous speed) and it is negative for positive slip (speed lower than synchronous speed). For super-synchronous speed operation, P_r is transmitted to DC bus capacitor and tends to raise the DC voltage. The design and performance evaluation of variable speed wind generation system [3]. For sub-synchronous speed operation, P_r is taken out of DC bus capacitor and tends to decrease the DC voltage. C_{grid} is used to generate or absorb the power P_{gc} in order to keep the DC voltage constant. In steady-state for a lossless AC/DC/AC converter P_{gc} is equal to P_r and the speed of the wind turbine is determined by the power P_r absorbed or generated by C_{rotor} .

2.2 Indirect Vector Control:

The Figure 3 explains the fundamental principle of indirect vector control with the help of a phasor diagram. The d^s - q^s axes are fixed on the stator but the d^r - q^r axes, which are fixed on the rotor, are moving at speed ω_r as shown. Synchronously rotating axes d^e - q^e [4] are rotating ahead of the d^r - q^r axes by the positive slip angle Θ_{sl} corresponding to slip frequency ω_{sl} . Since the rotor pole is directed on the d^e axes and $\omega_e = \omega_r + \omega_{sl}$ ----- (6)

we can write

$$\Theta_e = \int \omega_e dt = \int (\omega_r + \omega_{sl}) dt = \Theta_r + \Theta_{sl} \text{----- (7)}$$

The rotor pole position is not absolute, but is slipping with respect to the rotor at frequency ω_{sl} . The phasor diagram suggests that the decoupling control, the stator flux component of current i_{qs} should be on the q^e axis, as shown

For decoupling control, we can now make a derivation of control equations of indirect vector control with the help of d^e - q^e equivalent. The rotor circuit equation can be written as

$$\frac{d\psi_{dr}}{dt} + R_r i_{dr} - (\omega_e - \omega_r) \psi_{qr} = 0 \text{----- (8)}$$

$$\frac{d\psi_{qr}}{dt} + R_r i_{qr} - (\omega_e - \omega_r) \psi_{dr} = 0 \text{----- (9)}$$

The rotor flux linkage expression can be given as

$$\psi_{dr} = L_r i_{dr} + L_m i_{ds} \text{----- (10)}$$

$$\psi_{qr} = L_r i_{qr} + L_m i_{qs} \text{----- (11)}$$

From the above equations, we can write i_{dr} and i_{qr} as

$$i_{dr} = \frac{1}{L_r} \psi_{dr} - \frac{L_m}{L_r} i_{ds} \quad \& \quad i_{qr} = \frac{1}{L_r} \psi_{qr} - \frac{L_m}{L_r} i_{qs} \text{----- (12)}$$

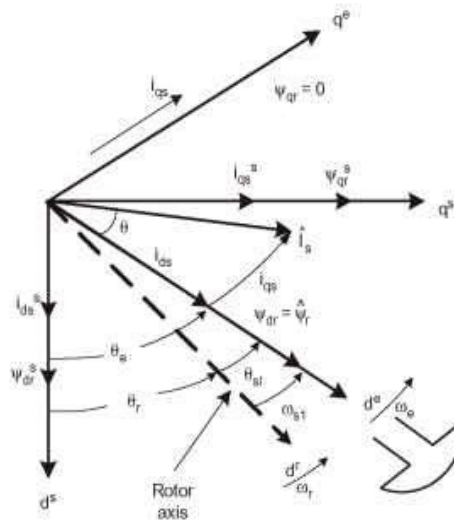


Fig 3: phasor diagram of indirect vector control

The rotor current in above equations which are inaccessible, can be eliminated with the help of equations of i_{dr} and i_{qr} as

$$\frac{d\psi_{dr}}{dt} + \frac{R_r}{L_r} \psi_{dr} - \frac{L_m}{L_r} R_r i_{ds} - w_{sl} \psi_{qr} = 0 \text{---- (13)}$$

$$\frac{d\psi_{qr}}{dt} + \frac{R_r}{L_r} \psi_{qr} - \frac{L_m}{L_r} R_r i_{qs} + w_{sl} \psi_{dr} = 0 \text{---- (14)}$$

Where $w_{sl} = w_e - w_r$ has been substituted.

For decoupling control, it is desirable that

$$\psi_{qr} = 0$$

That is
$$\frac{d\psi_{qr}}{dt} = 0 \text{--- (15)}$$

So that the total rotor flux ψ_r is directed on the d^e axis. By substituting the above equations we get

$$w_{sl} = \frac{L_m R_r}{\psi_r L_r} i_{qs} \text{----- (16)}$$

The frequency signal can be estimated as follows $\cos\theta_e = \psi_{ds}^s / \psi_s$ and $\sin\theta_e = \psi_{qs}^s / \psi_s$ ---- (17)

III. POWER CIRCUIT AND CONTROL STRATEGY

The Turbine at the left (a vertical type) is coupled to the cage – type induction generator through a speed up gear ratio. The variable frequency, variable voltage power generated by the machine is rectified to dc by a PWM voltage fed rectifier that also supplies the excitation current (lagging) to the machine.



3.1 Generator Speed Tracking Control (FLC-1)

Fig 4: Block Diagram of FLC -1

If on the other hand $+\Delta w_r$ causes $-\Delta P_0$, the direction of search is reversed. The variables ΔP_0 , Δw_r and $L\Delta w_r$ are described by membership functions and rule table. In the implementation of fuzzy control, the input variables are fuzzified, the valid control rules are evaluated and combined, and finally the output is defuzzified to convert to the crispy value. The wind vortex and torque ripple can lead the search to be trapped in a minimum which is not global, so the output Δw_r is added to some amount of $L\Delta w_r$ in order to give some momentum to continue the search and to avoid such local minima. The controller operates on a per-unit basis so that the response is insensitive to system variables and the algorithm is universal to any system. The membership functions of fuzzy logic controllers are explained in [4]. The scale factors KPO and KWR, as shown in Fig. 4, are generated as a function of generator speed so that the control becomes somewhat insensitive to speed variation. The scale factor expressions are given, respectively, as

$$KPO = a_1 w_r$$

$$KWR = a_2 w_r$$

Where a_1 and a_2 are the constant coefficients that are derived from simulation studies. Such coefficients are converting the speed and power in per-unit values. The advantages of fuzzy control are obvious. It provides adaptive step size in the search that leads to fast convergence, and the controller can accept inaccurate and noisy signals. The FLC-1 [5] operation does not need any wind velocity information, and its real time based search is insensitive to system parameter variation

3.2 Generator Flux Programming Control (FLC-2)

Since most of the time the generator is running at light load, the machine rotor flux can be reduced from the rated value to reduce the core loss and thereby increase the machine-converter system efficiency. The principle of online search based flux programming control by a second fuzzy controller FLC-2 is explained in Fig.5. This causes increasing torque current i_{qs} by the speed loop for the same developed torque. As the flux is decreased, the machine iron loss decreases with the attendant increase of copper loss. However, the total system (converters and machine) loss decreases, resulting in an increase of total generated power P_0 .

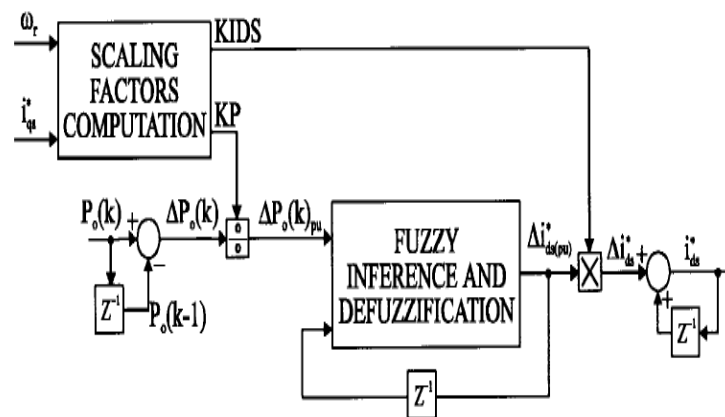


Fig 5: Block diagram of FLC-2

The principle of fuzzy controller FLC-2 is somewhat similar to that of FLC-1. The system output power $P_0(k)$ is sampled and compared with the previous value to determine the increment ΔP_0 . In addition, the last excitation current decrement ($L\Delta i_{ds}$) is reviewed. On these bases, the decrement step of i_{ds} is generated from fuzzy rules through fuzzy inference and defuzzification, as indicated. It is necessary to process the inputs of FLC-2 in per-unit values. Therefore, the adjustable gains KP and KIDS convert the actual variable to variables with the following expressions

$$KP = a w_r + b$$

$$KIDS = c_1 w_r - c_2 i_{qs} + c_3$$

where a , b , c_1, c_2 and c_3 are derived from simulation studies. The current i_{qs} is proportional to the generator torque, and $\Delta\omega_r$ is zero because the fuzzy controller FLC-2 is exercised only at steady-state conditions. The FLC-2 controller operation starts when FLC-1 has completed its search at the rated flux condition. If wind velocity changes during or at the end of FLC-2, its operation is abandoned, the rated flux is established, and FLC-1 control is activated.

3.3 Closed-Loop Generator Speed Control (FLC-3)

The speed loop control is provided by fuzzy controller FLC-3, as indicated in Fig. 6. As mentioned before, it basically provides robust speed control against wind vortex and turbine oscillatory torques. The disturbance torque on the machine shaft is inversely modulated with the developed torque to attenuate the modulation of output power and prevent any possible mechanical resonance effect. In addition, the speed control loop provides a deadbeat type response when an increment of speed is commanded by FLC-1.

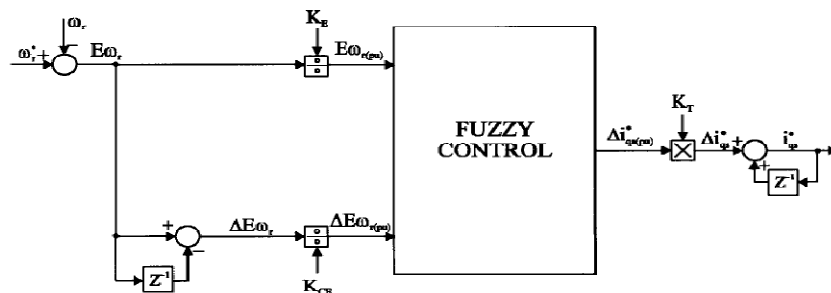


Fig 6: Block Diagram of FLC - 3

The speed loop error ($E\omega_r$) and error change ($\Delta E\omega_r$) signals are converted to per-unit signals, processed through fuzzy control, and then summed to produce the generator torque component of current. Note that, while fuzzy controllers FLC-1 and FLC-2 operate in sequence at steady (or small turbulence) wind velocity, FLC-3 is always active during system operation.

IV. SIMULATION RESULTS

Wind - generation system is simulated to validate all the control strategies and then evaluate the performance of the system.

The machine and turbine parameters are given as:

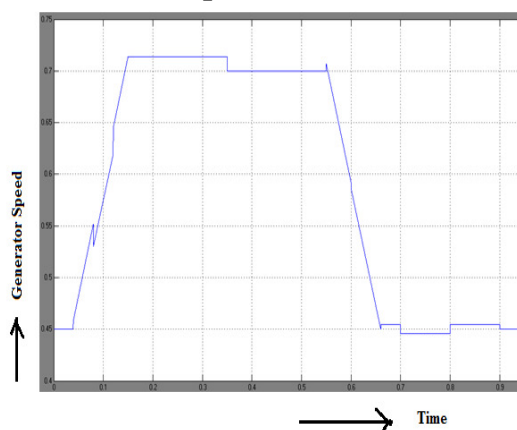
Machine Parameters:

3 Phase, 7 hp, 230/450v, 7.6A, 4 poles,
1500rpm, $R_l = 0.25\text{ohm}$, $R_r = 0.335\text{ohm}$

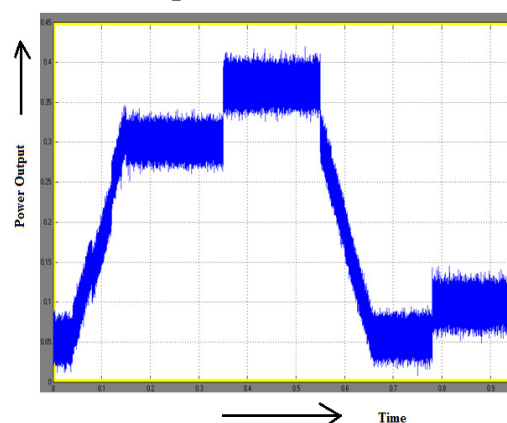
Turbine Parameters

3.5KW, Tower = 99.95m, 11.1-22.2r.p.m, $\eta_{\text{gear}} = 5.2$,
 $A = 0.015$, $B = 0.03$, $C = 0.015$

Generator Speed ~ Time



Output Power ~ Time



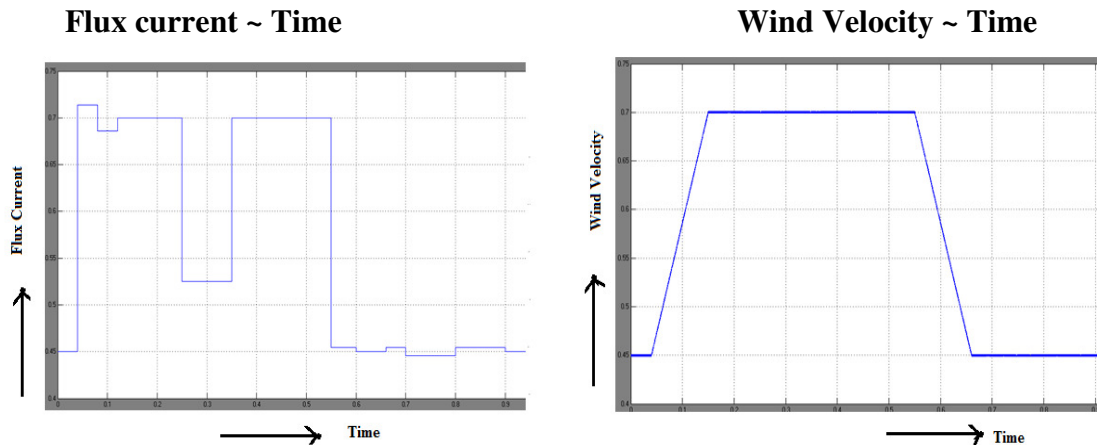


Fig 7. Simulation Results

Simulation of wind generation system is performed in matlab and results are presented in fig 7. Generator speed, output power, flux current and wind velocity with respect to time are plotted.

V. CONCLUSION

The fuzzy logic based variable speed cage machine wind generation system has been analyzed. The system performances have been studied with matlab- simulation to validate all the theoretical concepts. There are three fuzzy logic controllers in the generation system:

- The first fuzzy controller FLC-1 searches on line the optimum generator speed so that the aerodynamic efficiency of the wind turbine is maximum.
- The second fuzzy controller FLC-2 programs the machine flux by an on line search so as to optimize the machine converter efficiency.
- The third fuzzy controller FLC-3 performs robust speed control against turbine oscillatory torque and wind vortex.

The main conclusions of this paper are:

- The system was found to be parameter insensitive with fuzzy controllers.
- The system shows a fast-convergence with fuzzy controllers.
- The system can accept noisy and inaccurate signals.
- The fuzzy algorithms used in the system are universal and can be applied retroactively in any other system.
- The performance of the system was found to be excellent with all the fuzzy logic controllers.

REFERENCES

- [1] K.Kaur,Dr.S.Chowdhury,Dr.S.P.Chowdhury,Dr.K.B.Mohanty,Prof.A.Domijan“Fuzzy Logic Based control of variable speed Induction machine wind generation ststem” IEEE Transactions.
- [2] Simoes , M. G. , Bose, B. K. , Spiegel , R.J. “Fuzzy logic based intelligent control of variable speed cage wind generation system” IEEE Transactions on Power Electronics Vol. 12 , No.1, (January 1997):pp.87-95 .
- [3] Simoes , M. G. , Bose , B. K. , Spiegel , R.J. “Design and performance evaluation of fuzzy logic based variable speed wind generation system”, IEEE Transactions on Industry Applications. Vol. 33 ,No. 4 , (July/august 1997), pp. 956-965.
- [4] Bose, B.K. Modern power electronics and A.C drives pp368-378
- [5] Zhao, Jin. , Bose , B.K. “Evaluation of membership functions for fuzzy logic controlled induction motor drive”, IEEE Transactions on Power Electronics , (2002), pp. 229-234.
- [6] C.C.Lee , “Fuzzy logic in control system- Fuzzy Logic controllers –I” , IEEE Transactions on Systems, Man and cybernetia,20(2),pp.404-418 , 1990
- [7] Bose, B.K. Modern power electronics and A.C drives, 2002.page

- [8] Souusa, G.C.D. , B.K.Bose , B.K., J.G.Cleand J.G."Fuzzy logic based on-line efficiency optimization control of an indirect vector – controlled induction motor drive", IEEE Transactions on Industrial Electronics,(April-1995),Vol. 42, No. 2.
- [9] Bhadra ,S.N. , Kastha , D. , Banerjee , S. Wind electrical system .New Delhi : Oxford Education ,2005.

Authors

B. Ravichandra Rao has received B.Tech from Sri Krishnadevaraya University, Anantapur in 2002 and M.E from Pune University,Pune in 2004 and pursuing Ph.D in electrical engineering from S.V.University, Tirupathi. He is presently working as Assistant Professor of EEE Department, G.Narayanamma Institute of Technology and Science, Hyderabad, India.



Amala Lolly.R received B.Tech from Jawaharlal Nehru Technological University Kakinada in the year 2009 and pursuing M.Tech in G.Narayanamma Institute of Technology and Science (for women), Hyderabad, India.



AN ADVANCED WIRELESS SENSOR NETWORK FOR LANDSLIDE DETECTION

Romen Kumar.M¹ & Hemalatha²

¹Research Scholar, Department of Computer Science

²Head, Department of Software System
Karpagam University, Coimbatore-21, India

ABSTRACT

The power of wireless sensor network technology has provided the capability of developing large scale systems for real-time monitoring. This paper describes the evolution and generation of a wireless sensor network system for landslide detection in north eastern state of India, a region known for its heavy rainfall, steep slopes and frequent landslides. The post to different places and data retrieval or data collection from geographical sensors, the design, deployment and deployment of data collection and data aggregation algorithms needed for the network, and the network requirements of the deployed landslide detection system, data analysis system etc has been discussed in this paper.

KEYWORDS: *Wireless Sensor Networks, Distributed Aggregation Algorithms, Heterogeneous Networks, Landslide.*

I. INTRODUCTION

The real-time monitoring of environmental disasters are one of the most the importance necessity of the world. Different type of technologies has been developed for this purpose. Wireless sensor network (WSN) is one of the major technologies that can be used for real-time monitoring. WSN has the capability of large scale deployment, low maintenance, scalability, adaptability for different scenarios etc. WSN has its own restriction such as low memory power and bandwidth etc, but its capability to be deployed towards environment, and low maintenance requirement made it one of the best suited technologies for real time monitoring.

This paper discusses the design and development of a land slide dictation system using WSN at north eastern state of India. The deployment site has historically experienced several landslides, with the latest one occurring in the year 2008, this remainder of the paper is organized as follows. Section II describes the methods for landslide prediction related work in WSN systems, and other related work in WSN systems. In Section III, we describe about landslide phenomena and Section IV describes about the sensors needed for monitoring rainfall induced landslides. Section V details about the enhanced sensor column design used along with this system. Section VI describes about the wireless sensor architecture used for landslide scenario and Section VII details the different WSN algorithms implemented in the landslide detection network. The wireless sensor testbed is described in details in Section VIII Field deployment; its design concerns and experience are described in Section IX. Finally we conclude in Section X and in same section future work is also discussed

II. RELATED WORKS

The evolution of WSN ha encourages the development of real-time monitoring of critical and emergency application. The wireless sensor technology has generated enthusiasm in computer scientists to learn and understand other domain areas which help to propose or develop real-time

deployments. The major area of focus is environment monitoring, detection and prediction. The Drought Forecast and Alert Systems (DFAS) has been proposed and developed in paper [1, 10]. This system uses mobile communication to alert the users, whereas the deployed system uses real time data collection and transmission using the wireless sensor nodes, Wi-Fi, satellite network and also through internet. The real streaming of data through broadband connectivity provides connectivity to wider audience.

An experimental soil monitoring network using a WSN is presented in reference [3] which explores real-time measurements at temporal and spatial granularities which were previously impossible. This paper also discusses about the reception of real-time measurements at temporal and spatial granularity. Research has shown that other than geotechnical sensor deployment and monitoring, other techniques such as remote sensing, automated terrestrial surveys, and GPS technology etc also can be used by themselves or in combination with other technologies to provide information about land deformation [54]. Paper describes a state of the art system that combines multiple sensor type to provide measurements the perform deformation monitoring.

Reference the topics of slip surface localization in wireless sensor networks, which can be used for landslide prediction. A durable wireless sensor node has been developed [13] which can be employed in expandable WSN for remote monitoring of soil conditions in areas conducive to slope stability failures. This study incorporates both theoretical and practical knowledge from diverse domains such as landslides and geomechanics, wireless sensor, Wi-Fi and satellite networks, power saving solutions and electronic interface and design, which paved the design, development and deployment of a real time landslide detection system using a WSN [52].

III. LANDSLIDE

Landslide is general term used to describe the down slope movement of soil, rock and organic materials under the influence of gravity. It can be triggered by gradual processes such as weathering or by external mechanism including:

- Undercutting of a slope by stream erosion, wave action, glacier, or human activity such as road building.
- Continuous rainfall, rapid snowmelt, or sharp fluctuation in ground-water levels.
- Shocks and vibration caused by earthquakes or construction activity.

Herby once the landslide is triggered, material is transport by various mechanisms including sliding, flowing and falling.

The types of landslides vary with respect to the:

- Rate of movement:- this range from a very slow creep (millimetres/year) to extremely rapid(meters/second)
- Type of material: - Landslide is composed of bedrock, unconsolidated sediment and organic debris.
- Nature of movement: - The moving remains can slide, slump, flow or fall.

Landslides constitute a major natural hazard in India that accounts for considerable loss of life and damage to communication routes, human settlements, agricultural fields and forest lands. The Indian continents with diverse physiographic, seism tectonic and climatologically conditions is subjected to varying degree of landslide hazards; the Himalayas including North-eastern mountains ranges being the worst affected, followed by a section of Western Ghats and the Vidhyas[2,3].

In India, landslides mainly happen due to heavy rainfall, so this study concentrates on rainfall includes landslide. Earthquakes can also cause landslides, however in India this is primary confined to the Himalayan region. High rainfall intensity accelerates the sliding and slumping in the existing hazard zones [44].

IV. SENSOR NEED FOR MONITORING RAINFALL INDUCED LANDSLIDES

Under heavy rainfall conditions, rain infiltration on the slope cause instability, a reduction in the factor of safety, pressure response, changes in water table height, a reduction in shear strength which holds the soil or rock, an increase in soil weight and a reduction in the angle of response. When the rainfall intensity is larger than the slope saturated hydraulic conductivity and runoff occurs.

There are three distinct physical events occur during a landslide:

- The preliminary slope failure
- The subsequent transport and
- The final evidence of the slide materials.

The initial slope failure can occur due to the increase in pressure and soil moisture content, under heavy rainfall which necessitates the inclusion of geophysical sensors for detecting the changes in pressure and moisture content with the warning system developed for landslide detection. So the system discussed in this paper also includes geophysical sensors such as pressure transducer and dielectric moisture sensor for capturing the measurement [5, 8].

After the slope failure the subsequent transport of the materials happens that will generate slope gradient change, vibration etc which has to be measured and monitored for effective issue of warning. So the warning system includes strain gauge and tilt meter that can be used for measuring slope gradient changes. Along with them geophone is used for analyzing the vibration [35, 42].

V. ENHANCED SENSOR COLUMN DESIGN

The commercially available wireless sensor node do not have implanted sensors to measure pressure, moisture content, vibration, earth movement etc. This constraint has led us to implement data acquisition boards to connect the external sensors to the wireless sensor nodes [60]. The geological sensors were placed inside a sensor column and they were connected to the wireless sensor node via a data acquisition board as shown in Figure 1. The sensor column design discussed is an enhanced version, which uses a homogenous structure, whereas our design uses a heterogeneous structure which differs with respect to the terrain conditions and geological and hydrological parameters of the deployment site [6]. And also in this sensor column design all the geological sensor such as geophone and dielectric moisture sensor are not placed inside the column but are connected to the same wireless sensor node. The sensor column design also includes tilt meters which can be used for validating the deformation measurements captured using strain gauges.

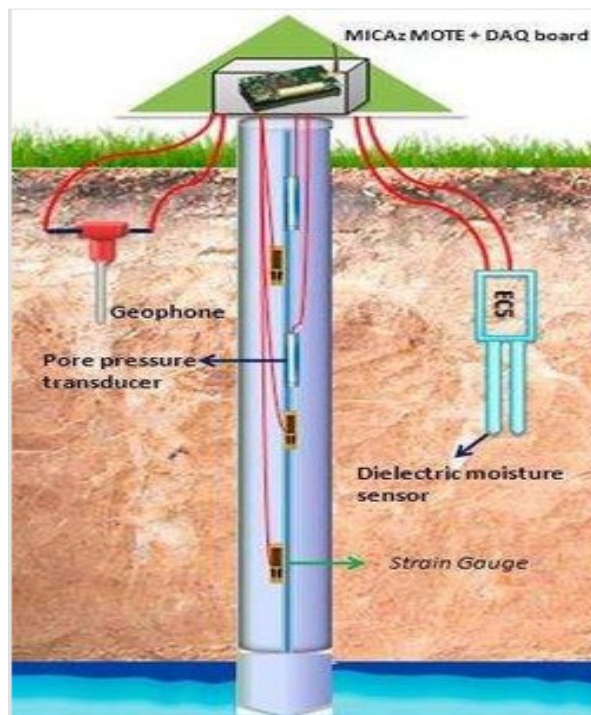


Figure1. Enhanced Sensor Column Design

VI. WIRELESS SENSOR NETWORK ARCHITECTURE

The WSN at the deployment site follows a two-layer hierarchy, with lower layer Wireless sensor nodes, sample and collect the heterogeneous data from the sensor column and the data packets are

transmitted to the upper layer. The upper layer aggregates the data and forwards it to the sink node gateway kept at the deployment site [28].

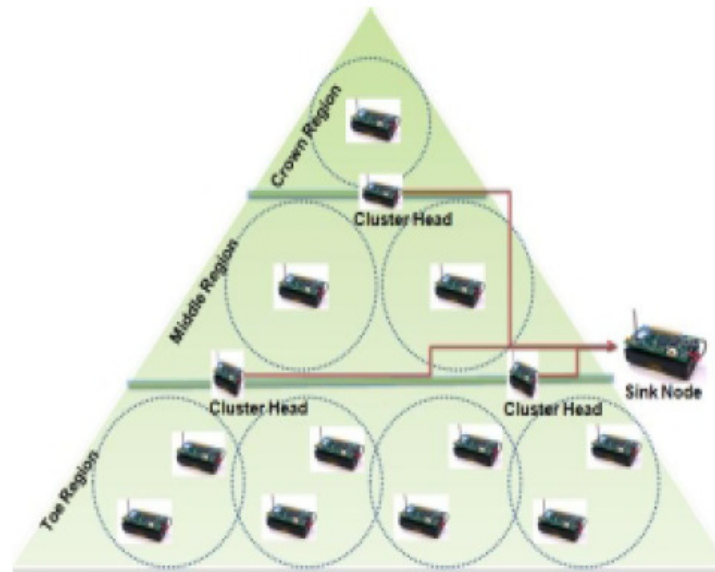


Figure 2.Regionalized WSN Architecture for Landslide

The geological and hydrological properties of the whole landslide area differ in each location, so it can be divided into a region having unique properties. Our deployment area is divided in three region such as crown region, middle region and toe region as shown in figure and numerous low level nodes attached to homogenous sensor column are deployed in these regions [7, 8].

VII. WIRELESS SENSOR NETWORK ALGORITHM

The WSN uses four algorithms for implementing clustering, distributed consensus among the data energy efficient data aggregation and time synchronization, which will contribute for the development of an efficient landslide detection system [25, 47].

The real-time monitoring networks are constrained by energy consumption, due to the remote location of the deployment site and the non availability of constant power. Considering this factor, the WSN at the deployment site implemented a totally innovative concept for distributed detection, estimation and consensus to arrive at reliable decisions, more accurate than that of each single sensor and capable to achieve globally optimal decisions as discussed in research papers [9, 22]. In landslide scenario, the implemented of this algorithm imposes a constraint of handling heterogeneous sensors in each sensor column. The different methods that can be used for implementing this algorithm, for landslide scenario are:-

- Homogenous sensor columns deployed in each region can be compared and a consensus value can be achieved for all the sensor columns in that region.
- All the sensors deployed in the landslide area can be assigned with a weightage with regards to its impact on landslide detection and a common consensus value can be achieved executing the algorithms at once, for all deployed sensors.
- Decentralized consensus performed for the same type of sensors in all sensor columns in a region.

Decentralized consensus for the same type of sensors has developed for the deployed network. The decentralized algorithms will be executed for each type of sensors, one by one, for all homogenous sensor columns deployed at each region. After initial set of sensor achieve its consensus, the next set of sensors will execute the decentralized algorithm and so on [30]. The other designs demand knowledge of correlation between different geophysical sensors, whereas this method does not require this knowledge, but the processing delay will be more compared to other methods, due to the multiple execution of same algorithm.

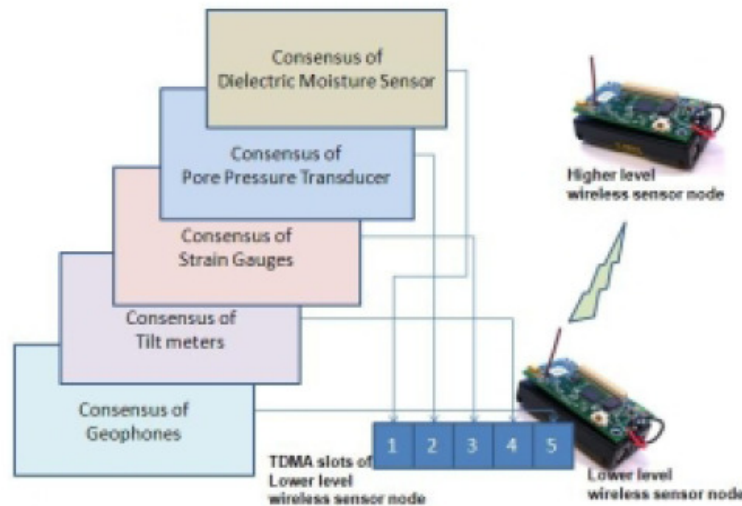


Figure 3. Decentralized Consensus for Same Type of Sensors

Since the study concentrates on the detection of rainfall induced landslide, the most relevant data will be arriving during rainy season. So rainfall based alert levels have been developed which will influence the sampling rate of the geological sensors and the transmission of data to higher layers as discussed in the threshold based algorithm [10, 11]. This algorithm will help to reduce the energy consumed during the low alert levels and also in collecting and transmitting large amount of data, only when the environment and geological conditions demand the same. Other than these methods, state level transitions have been incorporated to reduce the energy consumption per node which will also contribute to reduced energy consumption throughout the network [24, 32]. These requirements, however lead to the need of time synchronization and the algorithm planned for implementation in our network is discussed in research paper.

VIII. WIRELESS SENSOR TESTBED FOR LANDSLIDE DETECTION

The design and development of a WSN for the landslide scenario involves consideration of different factors such as terrain structure, vegetation index, climate, variation, accessibility of the area etc. The prerequisites of WSN development are selection of sensor column location, sensor column design and its data collection method, understanding transmission range and necessity of external antennas or addition relay nodes, identification of the communication protocol, development of application specific algorithms for data aggregation routing and fault tolerance etc [51].

The wireless sensor testbed deployed at northeast, India follow a two layer hierarchy, with the lower layer and an upper layer. The lower layer wireless sensor nodes are attached to the sensor column. They will sample and collect the heterogeneous data from the sensor column and the data packets are transmitted to the upper layer [12, 41]. The upper layer consists of cluster heads, which will aggregate the data and forwards it to the sink node gateway kept at the deployment site. Data received at the gateway has to be transmitted to the Field Management Center (FMC) which is approximately 500m away from the gateway. A Wi-Fi network is used between the gateway and FMC to establish the connection [39, 37].

The FMC incorporates facilities such as a VSAT (Very Small Aperture Terminal) satellite earth station and a broadband network for long distant data transmission. The VSAT satellite earth solution is used for data transmission from the field deployment site to the Data Management Centre (DMC) situated at our university campus 300 km away, while the broadband connection provides fault tolerance for long distance transmission and can be used for uploading real time directly to a web page within minimum delay [11, 20]. The DMC consist of the database server and an analysis station, which perform data analysis and landslide modelling and simulation on the field data to determine the landslide probability. The WSN architecture for land slide detection is shown in figure 4.

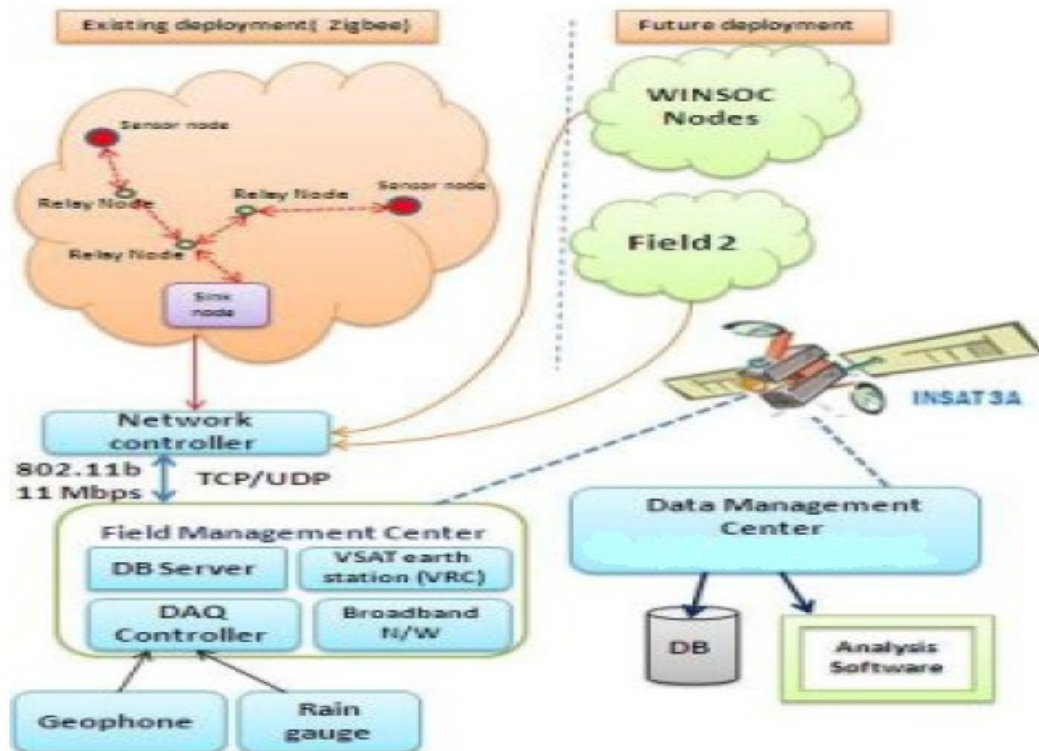


Figure 4.WSN Architecture for Landslide Detection

IX. FIELD DEPLOYMENT

The existing infrastructure has evolved through several iterative phases in its implemented. Important research focal points were deciding the sensor column locations, designing and constructing the sensor columns, sensor column deployment methods, interfacing circuitry, wireless sensor network, Wi-Fi network, satellite network, power solution, soil test and data analysis[59,20]. Extensive field investigations were conducted for identifying the possible locations for sensor column deployment. At the deployment site, an initial twenty sensor column location consisting of 150 sensors total, were identified with respect to their geological relevance [44]. The pilot deployment consist of two sensor columns, with ten sensors are deployed in the field along with six wireless sensor nodes as shown in figure 5.



Fig 5.Field Deployment

9.1. Deployment of Sensor Column

One of the sensor columns is deployed at the toe region where various water seepage lines converge. This facts lead to the installation of pressure transducer at different depth (2m, 5m) of the sensor

column 1[34]. The other geophysical sensors attached to its sensor column are dielectric moisture content transducer, and a geophone. Both pressure transducer and the dielectric moisture sensor are sampled at the rate of 10 samples/ second. The micaZ wireless sensor node connected to the sensor column transmits the digitized data values to the upper layers of the networks [13].

The other sensor column is attached with movement sensor since the location of it's in a unstable region. This sensor column has three tilt meters (91m, 2m, 3.5m) and three strain gauges (1.5m, 2.5m, and 4m) to capture the earth movement from the sensor column to the sensor column at 1 feet depth. The wireless sensor nodes sample three sensors at every five minutes and sent the data to upper level sensor nodes in the network [57, 26].

The spatial granularity will be increased by further addition of more sensor column approximately 20 and wireless sensor nodes approximately 20 which is in process.

9.2. Design and deployment of WSN

The design and development of WSN for the landslide scenario involves consideration of different factors such as terrain structure, vegetation index, climate variation, accessibility of the area etc. The prerequisites of WSN development are selection of sensor column location, sensor column design and its data collection method, understanding transmission range and necessity of external antennas or additional relay nodes, identification of the communication protocol, development of application specific algorithm for data aggregation, routing and fault tolerance etc[19,52]. The WSN architecture at the deployment is discussed section IV and the wireless sensor nodes used for the deployment are 2.4GHz MicaZ from Crossbow. The initial gateway was star gate with Intel XScale processor 500MHZ and running ARM Linux OS was programmed as the Sink Node while the new gateway is a single board computer which has 100MB RAM, 32 MB flash and a fixed base mote that is used to send and receive the messages through the transceiver [14].

The sensor column is physically attached to a wireless sensor node which is integrated with a data acquisition board. The distance between current sensor columns is approximately 50 meters, at a slope of about 70°. Due to the terrain structure and vegetation, the data for the sensor columns are not able to reach the gateway [23, 27]. The major reason for this is no line of sight path between the columns, between the first sensor column and the gateway, and between the second sensor column and the gateway. The observation along with experiment tests, have led us to employ three relay nodes in between the sensor columns themselves and the gateway. One of the relay nodes is a clustered head for this first and second column. The data from the cluster head is transmitted to the gateway in the form of packets. At the gateway the received packets are time stamped and stored [40, 46].

9.3. Deployment of Wi-Fi Network

The Wi-Fi network is used to transfer the data from the gateway to the FMC and it uses an external antenna and an access point for the same. The network has been tested with WLAN standards 802.11[31]. The Wi-Fi network allows us to install the gateway in any scalable distance from FMC. Since the region experiences frequent landslides and has several area within every 1 sqkm, which can be utilized as future extension sites for landslide detection systems by connecting them to FMC via a Wi-Fi networks [56,18].

9.4. Deployment of satellite Network

The basic satellite communication network in the landslide scenario is based on VSAT. The geological data collection at the landslide deployment site is transmitted from the FMC at the deployment site to the DMC, using the VSAT earth station [33]. The data is transmitted using UDP Protocol which includes recovery of lost packets, corrupted packets, secure transmission, and route via broadband during unavailability of VSAT, buffering the data to disk in case both networks are unavailable and sending the data as soon as the network is connected etc[48,29].

9.5. Monitoring Using Data Analysis System

The DMC consist of the database server and an analysis station which performs landslide modelling and data analysis on the data received from the field. The software also has the capability of real streaming of data and its analysis results, over internet which will provide greater capability of effective warning issue at minimum delay [45]. Data received at the DMC is being analyzed

using the in-house designed data visualization software which has the capability to determine factor of safety of the mountain and probability of landslide occurrence with respect to the signals received from the deployed sensors. It also has the capability to compare and analysis data from different sensor columns, different sensor in the same sensor column the same sensors in different sensor column, selective comparison etc [15, 17, 21,].

Data is successfully received from the deployment site with minimal data packet loss and analysis of data has been performed. Data received from two pressure transducers and a rain gauge is shown in Figure 6.

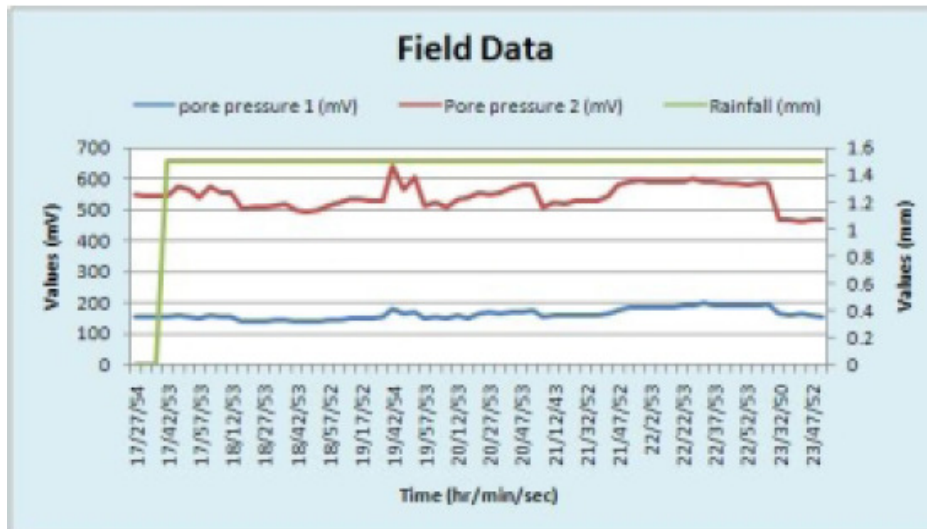


Fig 6.Real-time field data

During monsoon season, the sensors were able to capture the expansion and contraction of soil mass during heavy rainfall condition and after rainfall. The data analysis software showed respective variation in each of the deployed sensors.

X. CONCLUSION AND FUTURE WORK

WSN for landslide detection is one of the challenging research areas available today in the field of geophysical research. This paper describes about an actual field deployment of a WSN for landslide detection. This system uses a heterogeneous network composed of wireless sensor nodes, Wi-Fi and satellite terminals for efficient delivery of real time data to the data management canter. The data management canter is equipped with software's and hardware's needed for sophisticated analysis of the data. The results of the analysis in the form of landslide warnings and risks assessments will be provided to the inhabitant of the region.

In the future, this work will be extended to a full deployment with increased spatial variability, and the work in this regard is progressing. Field experiments will be conducted to determine the effects of density of the nodes, vegetation, location of sensor columns etc, for detecting rainfall induced landslide that may help in the development of low cost WSN for landslide detection.

ACKNOWLEDGMENT

We are very thankful to Karpagam University and Dr.M.Hemalatha for giving us advice and suggestions on the development of my research work.

REFERENCES

- [1]. Ramesh. M "Real time Wireless Sensor Network for Landslide Detection", proceedings of The Third International Conference on Sensor Technologies And Application, SENSORCOMN 2009,IEEE, Greece June 18-23, 2009.[2]
- [2]. Terzis Andreas, Anandarajah, Annalingam, Wang, I-Jing" Slip Surface Localization in Wireless sensor networks for Landslide Prediction, ISPN'06,USA, April 19-21, 2006.[9]

- [3]. Scutari.G,Barbarossa.S,” Distributed Consensus Over Wireless Sensor Networks Affected by Multipath Fading” IEEE Transaction on Signal Processing, Vol.56,No.8, August 2008.
- [4]. Garich.E “Wireless Automated Monitoring For potential Land Slide Hazards” Master Thesis, Texas A&M University, May 2007.[13]
- [5]. Raj.R, Kumar.S” Fault Tolerant Clustering Approaches in Wireless Sensor Network for Landslide Area Monitoring ” proceedings of the 2008 International Congress on Wireless Networks (ICWN’08), Vol.1, pages 107-113, CSERA Press, July 2008[15]
- [6]. Raiz Ahamed.S”The Effects of Global positioning System for Reliable Positioning, Navigation and Timing Services” Vol4No9/8Vol4No9, Journal of Theoretical and Applied Information Technology,2005
- [7]. Martinz.K, Hart.J.K “Environment Sensor Network”, IEEE Computer, Vol 37, 2004.
- [8]. Kung.H, Hau.J,Chen.C” Drought Forecast Model and Framework Using Wireless Sensor Networks” Journal of Information Science and Engineering, Vol 22, pages 751-769, 2006.
- [9]. Cogan,J, Szalay.A, Teriz.A “ A Wireless Soil Ecology Sensor Network”, 2006.
- [10]. Wang.G and Sassa.K “Pore Pressure Generation and Movement of rainfall-induced landslide: Effect of gain size and fine-particle content”, Engineering Geology Vol 36, pages 109-125,2003.
- [11]. Mckenna.G.T “Grouted In Installation of Piezometer in Boreholes” Canadian Geotechnical Journal, Vol 32, pages 353- 355, 1995.
- [12]. Marechal.M,Pierrot.J,Gorce.J “Fine synchronization for Wireless Sensor Networks using gossip averaging algorithms in proceedings of ICC 2008.
- [13]. Barbarossa.S, Scutari.G “Decentralized Maximum Likelihood Estimation for Sensor Networks Composed of Nonlinearly Coupled Dynamics Systems”, IEEE Transactions on Signal Proceeding, Vol 55, No 7, July 2007.
- [14]. Kunnath.A.T “Wireless Geophone for remote monitoring and Detection of landslides” International Conference, Communication and Signal Processing (ICCSP), 2011.
- [15]. Cao Jiong-qing”Hill Landslide Detection System Based on the IEEE 802.14 Wireless Sensor Networks”, Beijing Journal, Computer knowledge and Technology, 2008-2009.
- [16]. Arianna Pesci, Paolo Baldi, “ Digital Elevation Models for landslide evolution monitoring application on two areas located in Reno river valley (Italy)”, Annals of Geophysics, Vol 47, 2004.
- [17]. Bishwajeet Pradhan, Jasmi and Saolee “Probabilistic and Statistical Landslide hazard mapping using GIS and remote sensing at Cameron Highland, Malaysia”, 1999.
- [18]. Abbasi.I.A “Slope failure and landslide mechanism, Muree area, north Pakistan” Geological Bulletin University of Peshawar, Vol 35, 2003.
- [19]. Eric, James.S, Gardner “Modelling landslide hazards in the Kullu Valley, India using GIS and remote sensing”, Transactions on geosciences and remote sensing, IEEE, 2002.
- [20]. Yang Hong: An experiment global prediction system for rainfall-triggered landslide using satellite remote sensing and geospatial”, Transactions on geosciences and remote sensing, Vol.45, 2007.
- [21]. Hu.H, Fernandez, T.M.Dong, Azza,M “ LiDAR-based 3d FEM geological simulation and landslide stability analysis”, The 19th International Conference on Geoinformatics, China, June 24-26, 2011.
- [22]. Azzam.R, Fernandez, “Monitoring of landslide and infrastructure with sensor network in an earthquake environment”, The 5th International on Earthquake Geotechnical Engineering (5ICEGE), 2011.
- [23]. Fernandez-Steger, Rohn.J “Landslide after heavy rainfall: Causes and anthropogenic influence”, Workshop and fieldwork on Georisk, Eu- Project Asia link. Thailand, 2003
- [24]. Kempka.T, Waschusch.M, “ Influence of water content on underground”, Geoberlin,2006
- [25]. Lu, P., Stump, A., **Kerle, N.** and ... [et al.] Object - oriented change detection for landslide rapid mapping. In: IEEE Geosciences and remote sensing letters, 8 (2011)4 pp. 701-705, 2011.
- [26]. Stumps, A. and **Kerle, N.** Object - oriented mapping of landslides using random forests. In: Remote sensing of environment, 115 (2011)10 pp. 2564-2577, 2011.
- [27]. Wirawan, Ranchman S, Pratomo I, Mitta N. Design of Low Cost Wireless Sensor Networks - Based Environmental Monitoring System for Developing Country. Proc. Int. Conference APCC 14th Asia-Pacific. Tokyo, Japan. 2008.
- [28]. Jamal din MZ, Aripin NM, Isa AM, Mohamed HWL. Wireless Soil Temperature and Slope Inclination Sensors for Slope Monitoring System. Proceedings of International Conference on Energy and Environment (ICEE). Selangor, Malaysia. 2006.
- [29]. Jung woos Lee M S, Real-time Monitoring of Landslide using Wireless Sensor Network. PhD Thesis. Ohio: The Ohio State University; 2009.
- [30]. Herry Z Kotta, Silvester Tena, Gregarious Klau, K Rantelobo. Application of Geographical Information System (GIS) for Mapping Landslide Susceptibility: A Case Study of Timor Tengah Selatan, NTTProvince. Proceedings of National Seminar on Applied Technology, Science, and Arts (1st APTECS).Surabaya. 2009.

- [31]. Chuanhua Zhu, Xueping Wang. Landslide Susceptibility Mapping: A Comparison of Information and Weights-of-Evidence Methods in Three Gorges Area. International Conference on Environmental Science and Information Application Technology. Wuhan, China. 2009.
- [32]. Caine, N. (1980). The rainfall intensity-duration control of shallow landslides and debris flows, *Geografiska Annaler*, Vol. 62A, pp. 23-27, 1980.
- [33]. Garich, E. A. (2007). Wireless, Automated Monitoring for Potential Landslide Hazards, Master's Thesis, Texas A & M University, 2007. Iverson, R.M. (2000).
- [34]. Landslide triggering by rain infiltration, *Water Resource Research* Vol. 36, pp. 1897-1910, 2000. Kimura, H. & Yamaguchi, Y. (2000).
- [35]. Detection of landslide areas using satellite radar interferometer, *Photogrammetric Engineering & Remote Sensing*, Vol. 6 (3), pp. 337-344, 2000. Kumar, V.S.; Sampath, S.; Vinayak, P. V. S. S. K. & Harikumar, R. (2007).
- [36]. Rainfall intensity characteristics in at coastal and high altitude stations in Kerala, *Journal of Earth Systems Sciences*, Vol. 116 (5), pp. 451-463, 2007. Kung, H.; Hua, J. & Chen, C. (2006)
- [37]. Drought Forecast Model and Framework Using Wireless Sensor Networks, *Journal of Information Science and Engineering*, Vol. 22, pp. Kenneth, A. T. & Ramesh.M, M. V. (2010).
- [38]. Integrating geophone network to real-time wireless sensor network system for landslide detection, In *Proceedings of The Third International Conference on Sensor Technologies and Applications, SENSORCOMM 2010*, IEEE Explore, 2010.
- [39]. LAN, Hengxing. ZHOU, C.; Lee, C. F.; WANG, S. & Faquan, W. U. (2003). Rainfall induced landslide stability analysis in response to transient pore pressure - A case study of natural terrain landslide in Hong Kong, *Science in China Ser. E Technological Sciences*, Vol. 46, pp. 52-68, 2003.
- [40]. Liu, H.; Meng, Z. & Cui, S. (2007). A Wireless Sensor Network Prototype for Environmental Monitoring in Greenhouses, *IEEE Explore*, 2007. Wilson, R. C. & Wiczorek, G. F. (1995).
- [41]. Rainfall Thresholds for the Initiation of Debris Flow at La Honda, California, *Environmental and Engineering Geosciences*, Vol. 1 (1), pp. 11-27, 1995.
- [42]. Kyoji Sassa. Editors. *Landslide: Risk Analysis and Sustainable Disaster Management*. Proceeding of First General Assembly of the Int. Consortium on Landslide. Springer-Berlin. NY. 2005.
- [43]. WEN Hai-jia, LI Xin, ZHANG Jia-lan. An Evaluation-Management Information System of High Slope Geo-risk for Mountainous City Based on GIS. International Conference on Information Science and Engineering (ICISE). Nanjing, China. 2009; 1976-1978.
- [44]. Jung woos Lee M S, Real-time Monitoring of Landslide using Wireless Sensor Network. PhD Thesis. Ohio: The Ohio State University; 2009.
- [45]. Cesarean Alippi, Romulo Camplani, Cristian Galperti, Manuel Roveri. Effective design of WSNs: from the lab to the real world. *Proceeding of IEEE International Conference on Sensing Technology (ICST)* 3rd. Taipei, Taiwan. 2008; 1-9
- [46]. K. Aberer, M. Hauswirth, and A. Salehi. Global Sensor Networks. *IEEE Communications Magazine*, special issue on Advances in Service Platform Technologies for Next Generation Mobile Systems, 2006.
- [47]. A. Arora, R. Ramnath, and E. Ertin. Exscal: Elements of an extreme scale wireless sensor network, 2005.
- [48]. G. Barrenetxea, O. Couach, F. Ingelrest, M. Krichane, K. Aberer, M. Parlange, and M. Vetterli. *SenSorscope: an Environmental Monitoring Network*. *Water Resources Research Journal*, 2008.
- [49]. D. Estrin, L. Girod, L. Pottie, and M. Srivastava. Instrumenting the world with wireless sensor networks. *Acoustics, Speech, and Signal Processing*, 2001. *Proceedings. (ICASSP'01)*. 2001 IEEE International Conference on, 4:2033-2036 vol.4, 2001.
- [50]. Straser, E.G. and Kiremidjian, A.S. A Modular, Wireless Damage Monitoring System for Structures, John A. Blume Earthquake Engineering Center Report No. 128, Stanford, CA, 1998
- [51]. Garich, E.A. and Blackburn, J.T. "Automated, Wireless Instrumentation for Monitoring of Potential Landslide Hazards." *Proc. 1st American Landslide Conf.*, Vail, CO, in press, 2007
- [52]. Aguado, L.E., O'Driscoll, C. Xia, P., Nurutdinov, K., Hill, C. and O'Breine, P. 2006. A Low-Cost, Low-Power Galileo/GPS Positioning System for Monitoring Landslides. *Navitec* http://www.ggphi.eu/monitoring_landslides.pdf. October 2006
- [53]. Aoi, S., Kanugi, T. and Fujiwara, H. "Trampoline effect in extreme ground motion". *Science*, Vol. 322, No. 5902, pp. 2008
- [54]. Arnhardt, C., Asch, K.; Azzam, R; Bill, R.; Fernandez-Steger, T.M.; Homfeld, S.D. ; Kallash, A.; Niemeyer, F; Ritter, H.; Toloczyki, M. and Walter, K. "Sensor based Landslide Early Warning System – SLEWS, 2007
- [55]. Garich, E. A. Wireless automated monitoring for potential landslide hazards. - Master Thesis; Texas A & M University, 48 pp, 2007

- [56]. IFRC “ World Disaster report ” - Focus on early warning, early action, International Federation of Red Cross and Red Crescent, 204 pp,2009
- [57]. Jibson, R.W., Harp, E.L. and Michael, J.A. “A method for producing digital probabilistic seismic landslide hazard maps”. Engineering Geology, Vol. 58, 2000.
- [58]. Munich Re “Topics Geo - Natural catastrophes 2008, Analyses, assessments, positions”. Knowledge Series, Number 302-06022, 50 pp, 2009
- [59]. Shou, K-J. and Wang, C-F. “Analysis of the Chiufengershan landslide triggered by the 1999 Chi-Chi earthquake in Taiwan”. Engineering Geology, Vol. 68, pp. 237–250, 2004
- [60]. Keefer, D.K. “Landslides caused by earthquakes”. Geological Society of America Bulletin, Vol. 95, pp. 406-421. Keefer, D.K. “Investigating landslides caused by earthquakes - A historical review”. Surveys in Geophysics, Vol. 23, pp. 473-510, 2002

Authors

M. Hemalatha completed MCA M.Phil., PhD in Computer Science and currently working as a Asst Professor and Head, dept of software systems in Karpagam University. Ten years of Experience in teaching and published Twenty eight papers in International Journals and also presented seventy-eight papers in various National conferences and one international conferences Area of research is Data mining, Software Engineering, bioinformatics, Neural Network. Also reviewer in several National and International journals



M. Romen Kumar is presently doing PhD in Karpagam University, under the guidance of Dr. M. Hemalatha HOD of Department of software systems, Coimbatore, Tamil Nadu, and India and has completed MCA degree in 2008 and B.C.A in 2005. Major research area is Advance Networking - Sensor Network. Three papers published in International Journal and one paper presented in International Conference.



EVALUATION OF PHONETIC MATCHING APPROACHES FOR HINDI AND MARATHI: INFORMATION RETRIEVAL

Sandeep Chaware¹ and Srikantha Rao²

¹Research Scholar, MPSTME, Mumbai, India

²Research Supervisor, MPSTME, Mumbai, India

ABSTRACT

In multilingual environment, the phonetic matching plays an important role in various aspects. Basically, the techniques for phonetic matching are useful for information retrieval when text is not clear or not interpreted fully. Irrespective of correct form of the keyword, the entered keywords for information retrieval should be matched phonetically and results should be displayed. Many approaches have been proposed for phonetic matching such as use of text-to-phonetic system in translator-based system, use of operators such as MLLike, code-based approaches or language-specific phonetic-rule based approaches etc. Each approach is having limitations. In this paper, we tried to find some of the limitations of using those existing approaches for Hindi and Marathi languages and proposed some solutions for phonetic matching used for information retrieval.

KEYWORDS: *Phonetic matching, text-to-phonetic, writing style, phonetic rules, threshold.*

I. INTRODUCTION

The rapidly accelerating trend of globalization of businesses and the success of e-Governance solutions require data to be stored and manipulated in many different natural languages. The primary data repositories for such applications need to be efficient with respect to multilingual data. Efficient storage and query processing of data spanning over multiple natural languages are of crucial importance in today's globalized world.

As our country is diversified by languages and approximately 10% of population is aware of English language, this diversity of languages is becoming a barrier to understand and acquainted in digital world. In order to remove the language barrier, information technology (IT) solutions can play a major role. A system should be developed and deployed with multilingual support so that it can serve all-regional community requirements [1]. However Government of India had already launched the program called Technology Development of Indian Languages (TDIL) under which there are many projects such as development of corpora, OCR, text-to-speech, machine translation, keyboard layouts and so on [2]. It has been found that when services are provided in native languages, it has been strongly accepted and used.

India is a multilingual country with 22 recognized languages and 11 written script forms [3] (In some literature the officially Indian recognized languages were 23 [4]). All the scripts are derived from Brahmi and order of alphabet is similar. They also share some characteristics like common phonetic based alphabet, non-linear and complex scripts, word order free, and no cases in Indian scripts. A very peculiar feature of Indian languages is that though vowels can occur independently at the beginning, they do not occur independently within a word or as the last character of a word [5].

India is a country with various linguistics people. In India, the language or script changes after every 20 kilometers approximately. Though English language is a global language, it cannot be used everywhere in India due to minimum percentage of literacy. We need native languages in order to reach rural population. There are many areas of applications where we have to keep the data in many

languages so that people can access those data in their native languages when they don't know English. For example, railway reservation system, state or central government schemes, sales tax records, income-tax records, land records etc. These records should be maintained in English or in native languages. English records will help us faster processing and analyzing, which helps to make decision in certain situation where as native language records will be useful especially for rural and uneducated people. From those records, they either will get information or can provide valid data if necessary so that further analysis may be possible.

The goal is to provide a seamless interface to the user crossing all the language barriers. It has been found that a user is likely to stay twice as long at a site and four-times more likely to buy a product or consume a service, if the information is presented in their native language. Today English on the web is down to 35% from 90% in 1995. The fraction of Internet users that are non-native English speakers has grown from about half in mid-90's to about two-third and it is predicted that the majority of information available in the Internet will be multilingual by 2012 [6].

In this paper, we have proposed possible solutions to handle Indian language issues related to phonetic solutions. We had proposed system which will handle syntactic issues similar to phonetic as information retrieval for Hindi and Marathi. The phonetic issues are being handled by developing a system which will work on phonetic rules for languages and should allow minor variations in pronunciation or writing style. This way Indian language issues can be handled with respect to input, conversion and display.

II. PHONETIC MATCHING ISSUES FOR HINDI AND MARATHI

There are many phonetic matching issues for Hindi and Marathi languages. Some have been described below and are addressed in the successive sections.

- If we consider on hand approaches proposed for English, there are many alphabets for which no codes have been assigned as per algorithms. So, we may face problems in using and interpreting those alphabets. For example, the alphabets or letters like ण, झ, छ, and घ does not have code to match in Hindi language.
- If someone misses or adds language alphabets to a string, the string will be either misinterpreted or the system will give wrong result.
- The pronunciation of people other than Hindi or Marathi language speaking community may vary. It will be of great challenge to interpret and process those strings and provide the information.
- Strings ending with vowels need to be handled separately.
- Also, the strings in Hindi may have ambiguity of using 'Matras' with vowels or consonants.
- Special characters like 'ऋ', use of Nukta and so on need to be handled differently.
- Verbal and visual equivalences between speech sounds (phonemes) and written sign (graphemes) need to be found out. Their relationships have to be found out.

So, we have to consider all issues mentioned above in order to match the strings phonetically in Hindi and Marathi languages. This section focuses on some of the issues those are not being handled in existing approaches.

III. FOUNDATIONS FOR PHONETIC MATCHING

3.1 Input Mechanism

There are various mechanisms provided to input the keyword in native languages especially in Indian languages. Some are described below. We have used Input Method Editor (IME) method because of simplicity.

- *Multilingual Physical Keyboards:* There are many multilingual physical keyboards available for inputting Indian languages, but they are not feasible because it increases the cost and most users don't have multilingual keyboards and so it would be a rigid approach.

- *Multilingual On-screen Keyboards:* They can be downloaded from the Internet. But for each language the user must be aware of the character mappings between the existing physical keyboard and onscreen keyboard.
- *Input Method Editor (IME):* Input Method Editor (IME) does transliteration. Transliteration is a mapping from one system of writing into another word-by-word or ideally letter-by-letter which is opposed to transcription that specifically maps the sounds of one language to the best matching script of another language [7].
- *Inscript Keyboard Layout:* This keyboard layout has been standardized in 1986 by DoE and addresses few concerns about languages. These concerns includes, first, people perceived Indian languages as very difficult to use on mechanical typewriters. There are difficulties in learning keyboard layout on vernacular typewriters. Second, there was no standardization on vernacular keyboard layouts [7]. Since our languages have a phonetic nature, this leads to the development of a common phonetic layout based on consonants and vowels alone. All compositions and conjuncts were now handled by a computer with intelligent algorithms. With this phonetic keyboard, one can work on multiple languages; it is easy to learn for infrequent users, is excellent for typists, and provides ease of use for Indian languages. Since it is common for all Indian scripts, it has been named as Inscript keyboard.
- *On-Screen Keyboard Layout with IME:* In our domain, we had considered the on-screen keyboard layout to input Hindi and Marathi language strings. The on-screen keyboard layouts for Hindi and Marathi are shown in appendix A. In order to use those on-screen keyboard layout, we have to download and install IMEs for Hindi and Marathi. We had downloaded from bhashaIndia.com website [7]. We added those two languages in the language bar of my computer from desktop. At right corner of the screen, a small icon with language options appeared. We can switch from one language to another by selecting the language from this icon.

3.2 Storage Mechanism

There are many multilingual database systems have been developed and deployed such as Oracle 9i, Microsoft SQL Server 2000, IBM DB2 Universal Server (7.0), and My SQL. Many support the encoding standards like Unicode, ISCII or NChar as data type. Some of encoding forms for those database systems are described below.

- *ASCII Encoding:* The original American Standard Code for Information Interchange (ASCII) code was a 7-bit code used to encode all characters of the English language and several special characters such as a dot or a semicolon. However, this original code did not encode the umlauts of some of the European languages. Thus, the ASCII code was extended by 1 bit (8-bit ASCII code) to encode these characters as well. ASCII codes represent the text used in computer, communication devices. It includes definitions of 128 characters as 33 non-printable control characters, 94 printable characters and space are considered as invisible graphic. The ASCII code is a subset of the Unicode [8].
- *ISCII Encoding:* Indian Script Code for Information Interchange (ISCII) is a coding scheme for representing various writing systems of Indian languages. It is a Unicode standard for Indian scripts [33]. ISCII uses 8 bit code which is an extension of the 7 bit ASCII code containing the basic alphabet required for the 10 Indian scripts which have originated from the Brahmi script [8]. The ISCII code table is a super set of all the characters required in the Brahmi based on Indian scripts. For convenience, the alphabet of the official script Devnagari has been used in the standard. This is described in detail in appendix B.
- *Unicode Encoding:* Unicode standard is the Universal character encoding standard, used for representation of text for computer processing. Unicode standard provides the capacity to encode all of the characters used for the written languages of the world. The Unicode standards provide information about the character and their use. This standard is very useful for computer users who deal with multilingual text, business people, linguists, researchers, scientists, mathematicians and technicians. It uses a 16 bit encoding that provides code point for more than 65000 characters (65536). It assigns each character a unique numeric value and name. The Unicode standard and ISO10646 standard provide an extension mechanism called UTF-16 that allows for encoding as many

as a million. Presently Unicode standard provides codes for 49194 characters. It is the default standard for multilingual data storage in any database system. Unicode is a uniform 2-byte encoding standard that allows storage of characters from any known alphabet or ideographic system irrespective of platform or programming environments. Unicode codes are arranged in character blocks, which encode contiguously the characters of a given script (usually single language) [11].

➤ Unicode or ISCII encoding uses separate code points for each character. Logical order is used in rendering rules which tends to correspond to pronunciation. They are supporting full consonant forms. For inputting the characters, they save lot of space, which increases memory efficiency [1].

➤ *The NChar data type:* SQL standard specifies a new data type as National Char, (referred to as NChar) large enough to store characters from any Indian language or script. We can use **nchar** when the sizes of the column data entries are probably going to be similar or **nvarchar** when the sizes of the column data entries are probably going to vary considerably. SQL-92 onward all standards support NChar data type for storing national characters.

3.3 Display Mechanism

We must consider two cases for displaying multilingual text. First, running a system from some terminal and second, running a system under a window system. Using terminal, a system just sends correctly encoded text to terminals and leaves the task of rendering multilingual text to them. The code conversion is done with accordance to a coding system specified for the system output. In a window system, a system takes responsibility of displaying multilingual text. Each character set is assigned the corresponding font. A collection of mappings from all character sets to the corresponding fonts is named fontset and is the basis for displaying each character. A fontset can be used according to the context. We had used the first approach since font may not be important. Each character is being displayed on the screen as a rendering form, where it is equivalent to its either Unicode value.

IV. PHONETIC MATCHING APPROACHES: EXISTING SYSTEMS

4.1 Translator-Based System

In this category, each string of a language is translated into a uniform representation by using a text-to-phonetic (TTP) system [9]. This system will translate each text string into phonetic form. This phonetic form is a set of an encoding standard IPA, in which all the alphabet characters are represented in phonetic form. For some of the Indian languages, either TTP systems may not be available or they need to be developed.

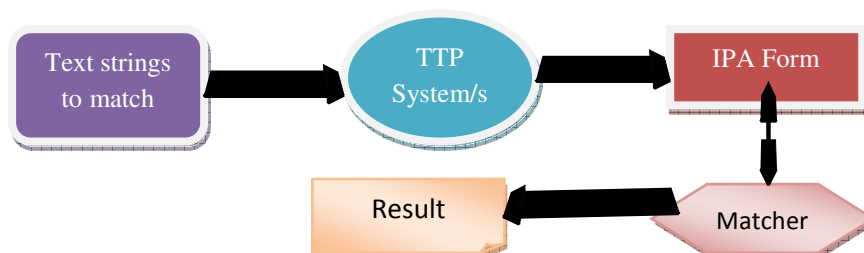


Figure 1. General Architecture of Text-To-Phonetic (TTP) Based System

Using this system, we cannot have the phonetic form of all the characters, especially for Hindi or Marathi. In order to match, edit distance can be calculated with some threshold value. Figure 3.1 shows the general architecture for text-to-phonetic based system.

4.2 Code-Based Systems

Using some code in the numeral form or other form, entire string is translated into a code format. This code always starts with first character of a string followed by at least four characters [10]. We need to compare the codes of both the strings to match two strings. If the codes are same then we may say that both strings are phonetically matching. The codes can be generated by grouping the alphabets according to their phonemes. Each group will have the same code value. Some system starts the code with 0, some starts with 1. But sometimes, we may get same code for different strings. Examples of

these systems are soundex, phonix and so on. Some systems may group the alphabet characters and assign the code for each group. If two strings are having maximum groups those are having same codes then we may say that both strings matches phonetically, example is Q-gram method.

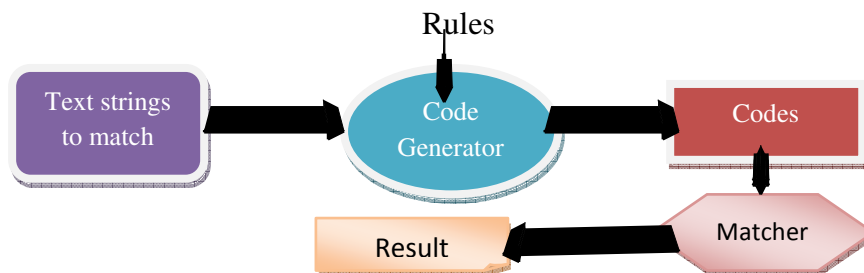


Figure 2. General Architecture of Code-based System

Figure 3.2 shows the general architecture for code-based system, where the codes will be generated by using rules. The matcher will match the codes for equivalence.

4.3 Phonetic-Rule Based Systems

These systems work on the phonetic rules designed for a particular language. These rules are used to group the alphabet characters according to phonemes. After applying these rules, each string is converted into its phonetic form either in text form or in some code form. In order to match, these forms are compared with some threshold value. These systems are easy to use, but difficult to build as we have to design phonetic rules for a language.

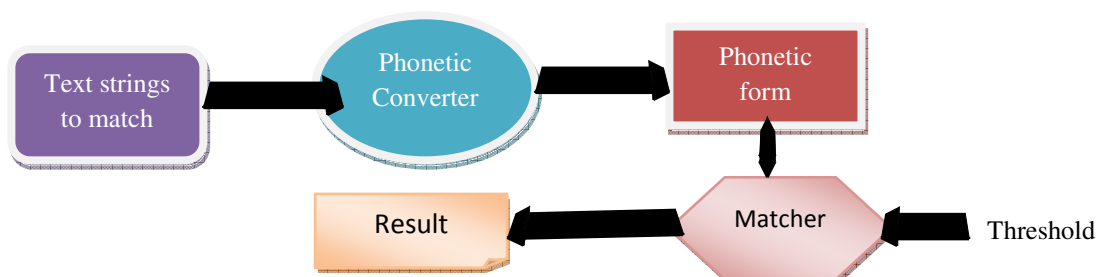


Figure 3. General Architecture of Phonetic-Rule based System

Figure 3.3 shows the general architecture for phonetic rule-based system, where the rules for each language apply to convert the string into its phonetic form. The matcher will use threshold value in order to match.

In this section, we elaborated the basic three approaches for phonetic matching. Those approaches may work for Hindi and Marathi languages, but need to be revised to a greater extent.

V. DRAWBACKS OF EXISTING PHONETIC MATCHING APPROACHES

The following are some of the drawbacks from existing phonetic approaches.

- In one of the approach, we need to find IPA code for each string for phonetic matching which is difficult and may not be available for Indian languages.
- Also we need to use text-to-phonetic (TTP) system for each language. Use of TTP makes the system complex.
- The algorithm depends on the user's predefined threshold value, so there may be an ambiguity in matching.
- The edit distance calculation is complex since many operations are to be carried out.
- The soundex and Q-gram methods use code for each alphabet. These methods are either generating wrong results or may not accommodate the code for all the alphabets for Hindi and Marathi languages.

VI. PROPOSED PHONETIC MATCHING APPROACHES

We proposed two phonetic approaches. One is based on writing style of the strings, where phonetic matching has been done by considering all possible writing styles of the native languages strings. Once matching has been done, information retrieval gives us the required results. In the second approach, we matched the strings phonetically by converting the strings into its equivalent phonetic form by using its phonetic rules for each language. These two matching approaches has been explained with proposed algorithm and example in the successive sections.

6.1 Phonetic Matching Approach - I

Objective: Phonetic Matching with Writing Style for Hindi and Marathi

Input: Native language string, S_{L1}

Output: IR in selected native language.

1. Enter the string in any native language such as Hindi or Marathi.
2. Parse the string to get vowels, consonants or modifiers.
3. Extract the vowels from the string.
4. Construct all possible combinations of string using vowels.
5. Convert the native language string into English by using mapping methodology.
6. Search the database based on all combinations.
7. Extract the result string from database.
8. Convert English language string/s from database into native language string/s.
9. Display the exact match in native language.

6.1.1 Example

Let's take an example of a Hindi string for which corresponding information has been retrieved.

String in Hindi: 'रघूलिला'

The following are the steps as per matching algorithm 5.2.3 to be applied to this string.

STEP 1 (Parsing): After taking native language string as input it is interpreted and parsed according to vowels, consonants and modifiers. Thus, we are getting a syllable for a string.

Parsing of a string 'रघूलिला': 'र ् अ घ ू ल ि ल ा'.

Figure 4. Parsing of a String 'रघूलिला'

The consonants are: र, घ, ल, ल

The vowels are: अ

The modifiers are: ्, ू, ि, ा.

Figure 4 shows the parsing of a string 'रघूलिला' as one of the possible ways of writing styles in Hindi or Marathi. Other possible ways of writing the same string are 'रघुलिला' or 'रघुलीला' or 'रघूलीला'. For each string, the system should matched phonetically and provide desired information. Similarly, we acquired the parsing of each string and used for matching.

Here, we are using full consonant approach so that we should get exact consonant, vowel or modifier. Even if we are using little high number of primitives for the string, it does not affect the inputting efficiency [1].

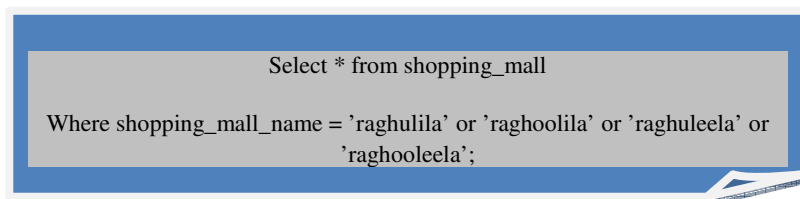
STEP 2 (Translation): Each native language string has to be translated into English, as we are maintaining the database in English for a shopping mall domain. We had used character-by-character mapping methodology for the translation. In this methodology, each character will be separately mapped as shown in table 1. This will convert native language string to English language string. After mapping, the entered Hindi string is translated to English as 'raghoolila' as per combinations of vowels, consonants and modifiers as shown in table 1.

Table 1. Hindi-To-English Conversion Mapping Table

Hindi Characters	र+्	अ	घ	ू	ल	ि	ल	ा
Equivalent English Characters	R	a	gh	oo	l	i	l	a
Equivalent ASCII Codes	2352	2309	2328	2370	2354	2311	2354	2310

STEP 3 (Query Formation): After conversion, a query is formed in SQL and fired against the database which is stored in English.

SQL Query: Select * from shopping_mall where shopping_mall_name = 'Raghoolila'. Similarly for all translated strings, SQL query is formed as in figure 5.

**Figure 5.** SQL Query for a String 'Raghulila' and Other Forms

The string is being passed to query module as a parameter and according to cases the query is formed. The string is searched in the corresponding database and retrieves it by the database module.

STEP 4 (Translation and Display): In order to convert English to native language string, we mapped each character with its ASCII code [7] and corresponding character is displayed, as shown in table 2. This task has been done by translation module.

Table 2. English-To-Hindi Conversion Mapping Table

Equivalent English Characters	R	a	gh	oo	L	i	l	a
Equivalent ASCII Codes	2352	2309	2328	2370	2354	2311	2354	2310
Equivalent Hindi Characters	र+्	अ	घ	ू	ल	ि	ल	ा

For the string 'Raghoolila', the entire tuple has been retrieved as information and translated into Hindi as per mapping methodology and shown as information.

6.1.2 Results

The results after phonetically matching as IR are shown in figures 6 and 7. Figure 6 shows the user interface to enter a string for phonetically matching. Figure 7 shows the IR result after phonetically matched with the existing database according to algorithm.

Figure 6. Sample Input Interface for Native Language

Figure 7. Result of Sample Query

6.2 Phonetic Matching Approach - II

Objective: Rule-based Phonetic Matching for Hindi or Marathi

Input: Two strings either in Hindi or Marathi to match OR one string for IR.

Output: Phonetic Matching Yes or No OR display of record/s from database as IR.

- Enter two strings Hindi or Marathi in order to match phonetically.
- Each string is translated into its phonetic form by using phonetic rules for each language.
- Parse those two strings to acquire combinations of vowels, consonants or modifiers.
- Obtain Unicode for each translated string by summing the Unicode value of each character of a string.
- Compare the resultant Unicode values of both the strings by considering a threshold value of 5%.
- If these values are within 5%, then we are saying that they are phonetically matched. Else they are not matching.
- For IR, the entered string is searched in database after converting into its equivalent phonetic form. If it matches by considering threshold value of 15%, then the corresponding tuple is displayed as IR.

6.2.1 EXAMPLE

Consider the two strings 'संतोष' and 'संथोष' in Hindi.

STEP 1 (Phonetic Equivalent Strings):

Its corresponding phonetic forms are:

संतोष = 'स॒अ॒न्अ॒त्ओष'

संथोष = 'स॒अ॒न्अ॒थ॒ओष'

STEP 2 (Parsing):

After parsing those two strings, we acquired the results as combinations of vowels, consonants and modifiers as:

स॒अ॒न्अ॒त्ओष = स॒ ॒ अ॒ न॒ ॒ अ॒ त॒ ॒ ओष

स॒अ॒न्अ॒थ॒ओष = स॒ ॒ अ॒ न॒ ॒ अ॒ थ॒ ॒ ओष

STEP 3 (Comparison):

After acquiring phonetic codes from Unicode of each character and transferring them to decimal values, we acquired the following codes for the strings:

संअनन्अत्ओष = 23487

संअनन्अथओष = 23488

By considering 5% threshold to match, the difference is calculated as:

$$(23488 - 23487) / 23488 * 100 = 0.0042\%$$

STEP 4 (Result):

The difference is within 5% threshold, so we can say that those strings are phonetically matched.

6.2.2 Results

Table 3 shows the comparison of various strings in Hindi and Marathi for phonetic matching. We compared our approach with soundex and Q-gram methods and obtained better and accurate results. The results are also shown in graphical form as in figure 8. Figure 9 and figure 10 shows the information retrieval results after phonetic matching as per proposed methodology.

Table 3: Comparison of Strings for Hindi and Marathi

Strings	HINDI			MARATHI		
	SOUNDEX	Q-GRAM	INDIC-PHONETIC	SOUNDEX	Q-GRAM	INDIC-PHONETIC
संतोष & संथोष	YES	YES	YES	YES	YES	YES
मुंबई & मुंवाई	YES	YES	NO	YES	YES	NO
रघुलिला & राघुलिला	YES	YES	YES	YES	YES	YES
संध्या & सैंडी	YES	YES	NO	YES	YES	NO

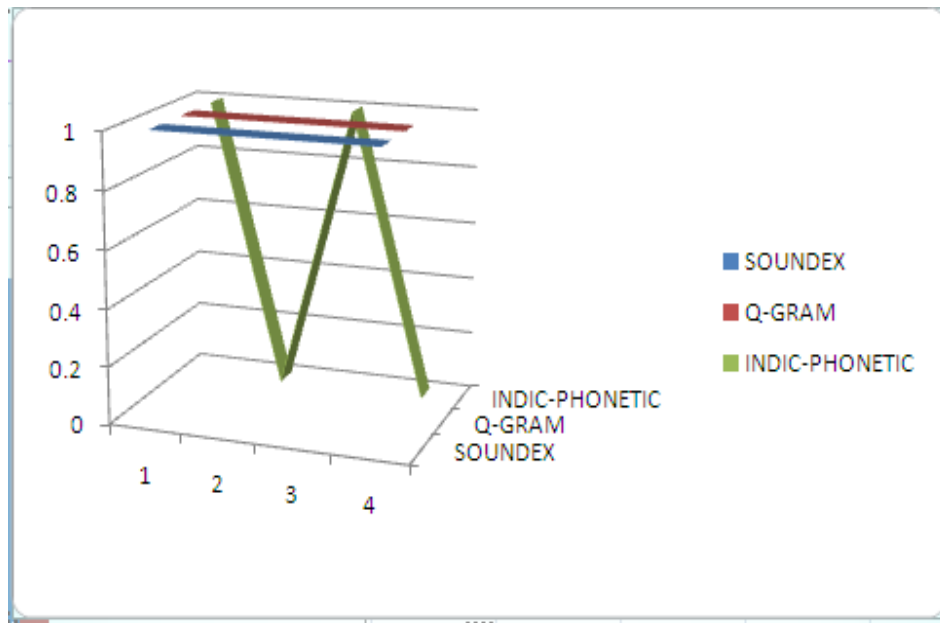


Figure 8: Graphical Comparison of Three Phonetic Matching Methods

Figure 9. Phonetic Name-wise Search in Domain for IR

Figure 10: IR after Phonetic Matching

VII. CONCLUSION

Many phonetic matching approaches, methods, algorithms have been proposed. But all these need lot of parameters, number of external resources needed for matching and so on. Basically all these methods are dependent on either international phonemic alphabet or translation system for each language. Some approaches rely on code for each alphabet or rules based on pronunciation for matching. In this paper, we classified the general approaches for phonetic matching. In proposed approaches, these classifications have been applied and evaluated. We also made an evaluation of our proposed approaches and compared with approaches like soundex, Q-gram which may work for English but may give wrong result for Hindi and Marathi languages. We found better and accurate results as compared to other existing approaches for our proposed approaches.

REFERENCES

- [1] Madhuresh Singhal et al. 'Developing Information Technology Solutions in Indian Languages: Pros and Cons'. Private Publication.
- [2] <http://www.tdil.mit.gov.in>
- [3] Ranbeer Makin et al. 'Approximate String Matching Techniques for Effective CLIR among Indian Languages'. Private Publication.
- [4] Pranav Mistry and Niranjana Nayak. 'AKSHAR: A mechanism for inputting Indic scripts on digital devices'. USID2007, June 18-20, 2007, Hyderabad, India.
- [5] Prof. R.K. Joshi et al. 'A Phonetic Code-based Scheme for Effective Processing of Indian Language'. Internationalization and Unicode Conference, Prague, Czech Republic, March 2003.
- [6] K. Ganesan and G. Siva. 'Multilingual Querying and Information Processing'. Information Technology Journal 6 (5), 2007, pp 751-755.
- [7] www.xs4all.nl/~wjsh/hindi/htm
- [8] www.bhashaindia.com

- [9] A. Kumaran 'Multilingual Information Processing on Relational Database Architectures'. PhD Thesis, IISC Bangalore, 2006.
- [10] Justin Zobel and Philip Dart. 'Phonetic String Matching: Lessons from Information Retrieval'. Private Publication.
- [11] www.unicode.org

Authors

Sandeep Chaware is a Research Scholar at MPSTME, NMIMS, Mumbai and his research area is 'Phonetic and Semantic matching Approaches for Hindi and Marathi'.



Srikantha Rao is a Director at TIMSCDR and Research supervisor at MPSTME, NMIMS, Mumbai.



DESIGN OF ENERGY-EFFICIENT FULL ADDER USING HYBRID-CMOS LOGIC STYLE

¹Mohammad Shamim Imtiaz, ²Md Abdul Aziz Suzon, ³Mahmudur Rahman

¹Part-Time Lecturer, Department of EEE, A.U.S.T, Dhaka, Bangladesh

²Part-Time Lecturer, Department of EEE, A.U.S.T, Dhaka, Bangladesh

³Ex- Student, Department of EEE, A.U.S.T, Dhaka, Bangladesh

ABSTRACT

We present new designs for full adder featuring hybrid-CMOS design style. The quest to achieve a good-drivability, noise-robustness and low energy operations guided our research to explore hybrid-CMOS style design. Hybrid-CMOS design style utilizes various CMOS logic style circuits to build new full adders with desired performance. We also classify hybrid-CMOS full adders into three broad categories based upon their structure. Using this categorization, many full adder designs can be conceived. The new full adder is based on XOR-XOR Hybrid CMOS model that gives XOR and XOR full swing output simultaneously. This circuit's outperforms its counterparts showing 4%-31% improvement in power dissipation and delay. The output stage also provides good driving capability and no buffer connection is needed between cascaded stages. During our experiments, we found out that many of the previously reported adders suffered from the problems of low swing and high noise when operated at low supply voltages. The proposed full adders are energy efficient and outperform several standard full adders without trading of driving capabilities and reliabilities. The new full-adder circuits successfully operate at low voltages with excellent signal integrity and driving capability. The new adders displayed better performance as compared to the standards full adder. The problem we face during the experiment leads us to different zones where efficient circuit can be developed using this new full adder.

KEYWORDS: Adders, Exclusive OR gate (XOR), Exclusive NOR gate (XNOR), Multiplexer, Hybrid-CMOS design style, low power.

I. INTRODUCTION

The necessity and popularity of portable electronics is driving designers to endeavor for smaller area, higher speeds, longer battery life and more reliability. Power and delay are the premium resources a designer tries to save when designing a system. The most fundamental units in various circuits such as compressors, comparators and parity checkers are full adders [1]. Enhancing the performance of the full adders can significantly affect the overall system performance. Figure 1 shows the power consumption breakdown in a modern day high performance microprocessor [2]. The data path consumes roughly 30% of the total power of the system [19] [23]. Adders are an extensively used component in data path and therefore careful design and analysis is required.

So far several logic styles have been used to design full adders. Each design has its own pros and cons. Classical designs use only one logic style for the whole full adder design. One example of such design is the standard static CMOS full adder [3]. The main drawback of static CMOS circuits is the existence of the PMOS block, because of its low mobility compared to the NMOS devices. Therefore, PMOS devices need to be sized up to attain the desired performance. Another conventional adder is the complementary pass-transistor logic (CPL) [3]. Due to the presence of lot of internal nodes and static inverters, there is large power dissipation. The dynamic CMOS logic provides a high speed of operation; however, it has several inherent problems such as charge sharing and lower noise immunity. Some other full adder designs include transmission-function full adder (TFA) [4] and

transmission –gate full adder (TGA) [5]. The main disadvantages of these logic styles are that they lack driving capability and when TGA and TFA are cascaded, their performance degraded significantly [23].

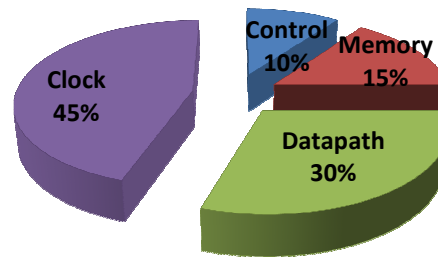


Figure 1: Power breakdown in high-performance microprocessors

The remaining adder designs use more than one logic style for their implementation which we call the hybrid-CMOS logic design style. Examples of adders built with this design style are DB cell [6], NEW 14-T adder [7], and hybrid pass logic with static CMOS output drive full adder [8] and new-HPSC [9] adder. All hybrid designs use the best available modules implemented using different logic styles or enhance the available modules in an attempt to build a low power full adder cell. Generally, the main focus in such attempts is to reduce the numbers of transistors in the adder cell and consequently reduce the number of power dissipating nodes. This is achieved by utilizing intrinsically low power consuming logic style TFA or TGA or pass transistors. In doing so, the designers often trade off other vital requirements such as driving capability, noise immunity and layout complexity. Most of these drivers lacking driving capabilities as the inputs are coupled to the outputs. Their performance as a single unit is good but when larger adders are built by cascading these single unit full adder cells, the performance degrades drastically [21] [25]. The problem can be solved by inserting buffers in between stages to enhance the delay characteristics. However, this leads to an extra overhead and the initial advantage of having a lesser number of transistors is lost.

A hybrid-CMOS full adder can be broken down into three modules [6]. Module-I comprises of either a XOR or XNOR circuits or both. This module produces intermediate signals that are passed onto Module-II and Module-III that generate Sum and C_{out} outputs, respectively. There are several circuits available in [1], [6] and [7] for each module and several studies have been conducted in the past using different combinations to obtain many adders [1], [6], [10].

This paper is structured as follows: Section 2 and its subsections briefly introduce three categorized model of full adder. Section 3 and its subsections represent our proposed circuits for three different Modules where we present a new improved circuit for the simultaneous generation of the XOR and XNOR outputs to be used in Module-I and propose a new output unit for Module-II and Module-III which consist of XOR-XNOR or Multiplexer. Using the new circuits in Module-I, II and III, we build new hybrid-CMOS full-adder cells which is discuss in Section 4. Section 5 briefly exhibits the results and discussion. The new adder is optimized for low power dissipation and delay then it is compared with the classical static-CMOS, CPL, TFA, TGA, NEW14T, HPSC, and NEW-HPSC full-adder cells. The proposed full-adder design exhibits full-swing operation and excellent driving capabilities without trading off area and reliability. Section 6 suggests the future work and modification of this paper. Section 7 concludes the paper.

II. FULL ADDER CATEGORIZATION

Depending upon their structure and logical expression we classified hybrid CMOS full adder cells [11] into three categories. The expression of sum and carry outputs of 1-b full adder based on binary input A, B, C_{in} are,

$$\text{Sum} = A \oplus B \oplus C_{in}$$

$$C_{out} = A.B + C_{in}(A \oplus B)$$

These output expression can be expressed in various logic style and that's why by implementing those logics different full adders can be conceived. Moreover, the availability of different modules, as

discussed earlier, provides the designer with more choices for adder implementation [21] [25]. Using these different modules [8] we suggest three possible structures for full adder and these are as follows.

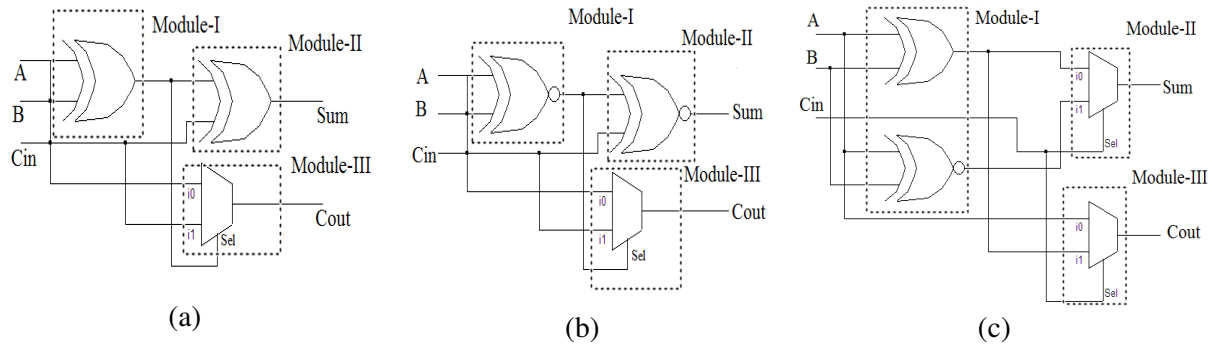


Figure 2: (a) General form of XOR-XOR based model, (b) General form of XNOR-XNOR based model, (c) General form of Centralized full adder

2.1 XOR-XOR BASED FULL ADDER

In this category, the Sum and Carry outputs are generated by the following expression, where H is equal to $A \oplus B$ and H' is the complement of H. The general form of this category is shown in Figure 2(a).

$$\text{Sum} = A \oplus B \oplus C_{in} = H \oplus C_{in}$$

$$C_{out} = A.H' + C_{in}.H$$

The output of the sum is generated by two consecutive two-input XOR gates and the C_{out} output is the output of a 2-to-1 multiplexer with the select lines coming from the output of Module-I. The Module-I can be either a XOR–XNOR circuit or just a XOR gate. In the first category, both Module-1 and Module-II consist of XOR gates. In the first case, the output of the XOR circuit is again XORED with the carry from the previous stage (C_{in}) in Module- II. The H and H' outputs are used as multiplexer select lines in Module-III. The Sum adders belonging to this category are presented in [12], [13].

2.2 XNOR-XNOR BASED FULL ADDER

In this category, the Sum and Carry outputs are generated by the following expression where A, B and C_{in} are XNORed twice to from the Sum and expression of C_{out} is as same as previous category. The general form of this category is shown in Figure 2 (b).

$$\text{Sum} = \overline{\overline{A \oplus B \oplus C_{in}}} = \overline{H' \oplus C_{in}}$$

$$C_{out} = A.H' + C_{in}.H$$

In this category, Module-I and Module-II consist of XNOR gates and Module-III consists of a 2-to-1 multiplexer. If the first module uses a XOR–XNOR circuit, then the H' output is XNORed with the C_{in} input to produce the Sum output. The static energy recovery full adder (SERF) [14] belongs to this category and uses a XNOR gate for Module-I and Module-II and a pass transistor multiplexer for Module- III.

2.3 CENTRALIZED FULL ADDER

In this category, the Sum and Carry outputs are generated by the following expression. The general form this category is shown in Figure 2(c).

$$\text{Sum} = H \oplus C_{in} = H.C'_{in} + H'.C_{in}$$

$$C_{out} = A.H' + C_{in}.H$$

Module-I is a XOR–XNOR circuit producing H and H' signals; Module-II and Module-III are 2-to-1 multiplexers with H and H' as select lines. The adder in [8] is an example of this category. It utilizes

the XOR–XNOR circuit presented in [7] and proposes a new circuit for output Module-III. The simultaneous generation of H and H' signal is critical in these types of adders as they drive the select lines of the multiplexers in the output stage. In another case (i.e. non simultaneous H and H'), there may be glitches and unnecessary power dissipation may occur. The final outputs cannot be generated until these intermediate signals are available from Module-I [20].

III. PROPOSED CIRCUIT FOR MODULE-I, II AND III

Hybrid CMOS full adders can be divided into three Modules. Each of the Models consists of XOR or XNOR or 2 to 1 multiplexer with selection lines. Module-I Consist of XOR or XNOR in all three categories; Module-II consists of XOR or XNOR for first two categories and 2 to 1 multiplexer for last category and Module –III consists of 2 to 1 multiplexer with selection lines in all three categories. Finally it can be said that three types of circuits used to form three categorized full adders. Here we will propose three new circuits for Module-I, Module-II and Module-III.

3.1 MODULE-I

Here we will talk about the proposed XOR and XNOR model. From the previous studies, we have found that XOR or XNOR gates based on transmission gate theory has limited transistor with enormous drawbacks. The drawbacks are the required complementary inputs and the loss of driving capability [14]. In general, if the output signals of a circuit come from V_{DD} or V_{SS} directly, we say this circuit has driving capability. If the circuit output will drive other circuits, it does better to cascade a canonical CMOS buffer to do so.

To follow without the loss of generality, all the methods we discuss will focus on the XOR function, mainly because the XNOR structure is very similar to XOR structure symmetrically. The skill for the XOR function can be applied to the XNOR function without question.

Based on the inverter configuration theory, two inverters are arranged to design XOR function as well as XNOR structure. These types of gates do not need the complementary signal inputs as like before and the driving property is better but it still have some defects such as no full driving capability on the output and more delay time [9].

In recent times simultaneous generation of XOR and XNOR has been widely used for Module-I, II [9], [14], [15]. This feature is highly desirable as non skewed outputs are generated that are used for driving the select lines of the multiplexer inside the full adder. Figure 3(a) shows a configuration using only six transistors and is presented in [14]. This circuit has been widely used to build full-adder cells [9], [14], [15]. The circuit has a feedback connection between XOR and XNOR function eliminating the non-full-swing operation [26]. The existence of V_{DD} and GND connections give good driving capability to the circuit and the elimination of direct connections between them avoids the short circuit currents component. However, when there is an input transition that leads to the input vector AB: XX-11 or AB: XX-00, there is a delay in switching the feedback transistors. This occurs because one of the feedback transistors is switched ON by a weak signal and the other signal is at high impedance state. This causes the increase in delay. As the supply voltage is scaled down, this delay tends to increase tremendously. This also causes the short circuit current to rise and causes the short circuit power dissipation to increase and eventually increase the power-delay product. To reduce this problem careful transistor sizing needs to be done to quickly switch the feedback transistors [9].

We found another improved version of XOR-XNOR circuit [8], [18], [26] which provides a full-swing operation and can operate at low voltages. The circuit is shown in figure 3(b). The first half of the circuit utilizes only NMOS pass transistors for the generation of the outputs. The cross-coupled PMOS transistors guarantee full-swing operation for all possible input combinations and reduce short-circuit power dissipation. The circuit is inherently fast due to the high mobility NMOS transistors and the fast differential stage of cross-coupled PMOS transistors. But the main drawback was it showed worse output at low voltage but at high voltage it showed completely opposite characteristic [18].

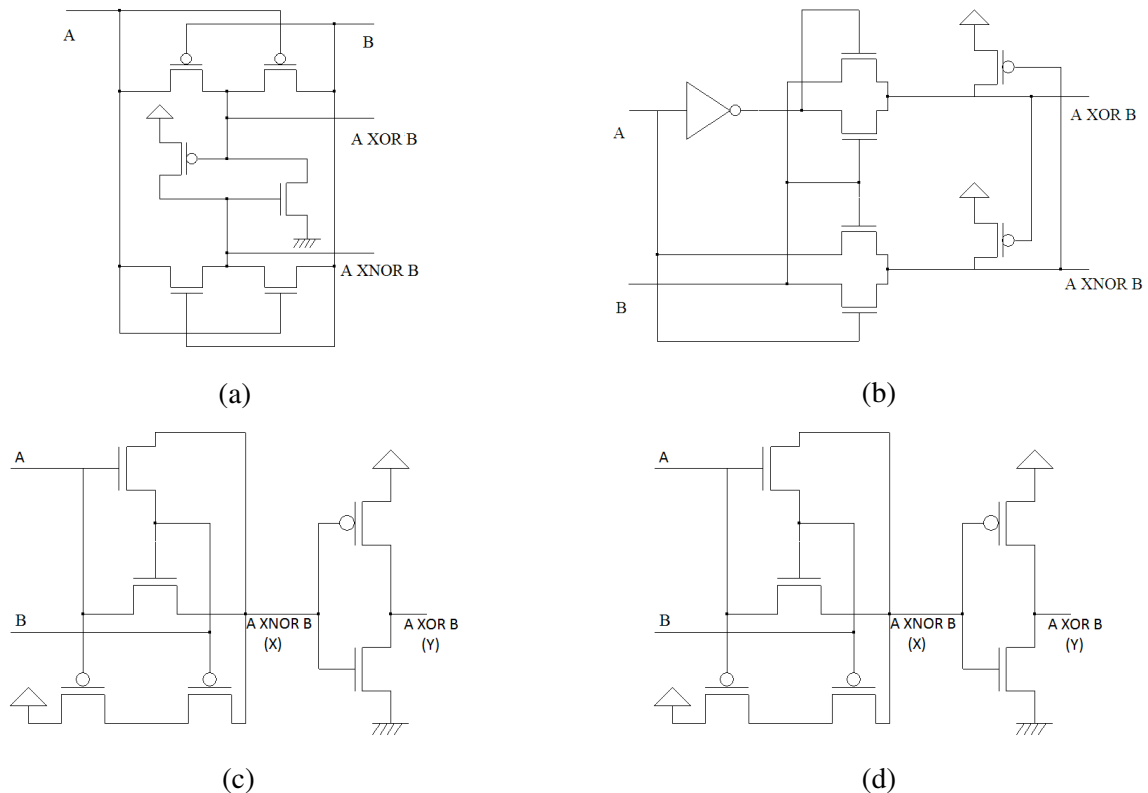


Figure 3: (a) Circuit 1 for XOR-XNOR model (b) circuit 2 for XOR XNOR model (C) Proposed XOR (d) proposed XNOR

We propose a novel XOR–XNOR circuit using six transistors that generates XOR and XNOR outputs simultaneously. Figure 3(c) and 3 (d) respectively represent Proposed XOR and XNOR circuit. On the 6-transistor design, the new proposed structures require non-complementary inputs and their output will be perfect. The initial plan was creating 4-transistor design but it was jeopardized due to worse output when both inputs were low for XOR and high for XNOR. Analysis of 4-transistor XOR structures, the output signal is the cases of input signal $AB=01, 10, 11$ will be complete. When $AB=00$, each PMOS will be on and will pass a poor low signal at the output end. That is, if $AB=00$ the output end will display a voltage, threshold voltage, a little higher than low but path driving capability exist, due to NMOS being on. Hence though the output is not complete, the driving current will increase. For XNOR function, the output in the case of $AB=00, 01, 10$ will be complete. While $AB=11$, each NMOS will be on and pass the poor high signal level to the output end. The analysis of driving capability is the same as XOR structure. By cascading a standard inverter to the XNOR circuit, a new type of 6-transistor XOR is found which will have a driving output, and the signal level at the output end will be perfect in all cases. The same property is presented in the 6-transistor XNOR structure. The proposed XOR-XNOR circuit was compared to circuits in figure 3(a) and 3(b) based on number of transistors, power and delay. In all the criteria our proposed model performs outstandingly. The simulation results at $2 V_{DD}$ and $2V$ input are shown in Table-I:

Table 1: Simulation results for the proposed XOR-XNOR Circuit at 50-MHz Frequency and $2V_{DD}$

	Circuit [1]	Circuit [2]	Propose XOR	Propose XNOR
No. of Transistor	6	10	6	6
Power (μW)	7.524	8.750	4.07	4.07
Delay (ns)	0.305	0.210	0.108	0.106

3.2 MODULE-II

Here we will review some of the existing and most frequently used circuits that can be used in the different modules of the full adder. From previous studies, we learned about eight different circuits [15], [16] which performed best in their available ways with advantages and disadvantages. Among eight of them we choose the best two and used the more efficient one for our proposed model. Those two circuits are given in figure 4

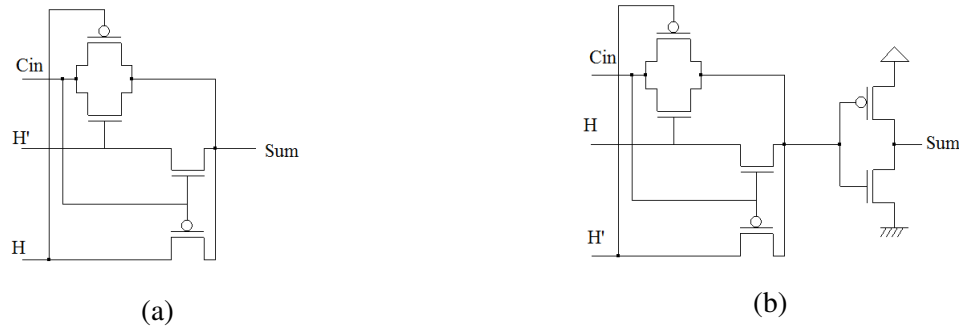


Figure 4: Circuits for Module-II

Figure 4(a) has transmission-function implementation of XOR and XNOR functions. This circuit does not have supply rails thereby eliminating short circuit current. Figure 4(b) is essentially the complement and has an inverter to produce Sum. This provides good driving capability due to the presence of the static inverter. This circuit is one of the best performers among all the circuits mentioned in [8] in terms of signal integrity and average power-delay product [6]. Both the circuits avoid the problem of threshold loss and have been widely used in adder implementation [15], [16]. We employ this circuit for our full-adder design.

3.3 MODULE-III

The expression of Module-III is,

$$C_{out} = A \cdot H' + C_{in} \cdot H$$

This expression is the output of 2 to 1 multiplexer with H and H' as the select lines. The most common implementation of the previous expression is using transmission gates (TG). Figure 5(a) shows the circuit for a 2-to-1 multiplexer using TG. The main drawback of this multiplexer is that it cannot provide the required driving capability to drive cascaded adder stages. One solution to this problem is to have an output buffer as shown in Fig. 5 (a). This would incur extra delay and an overhead of four transistors.

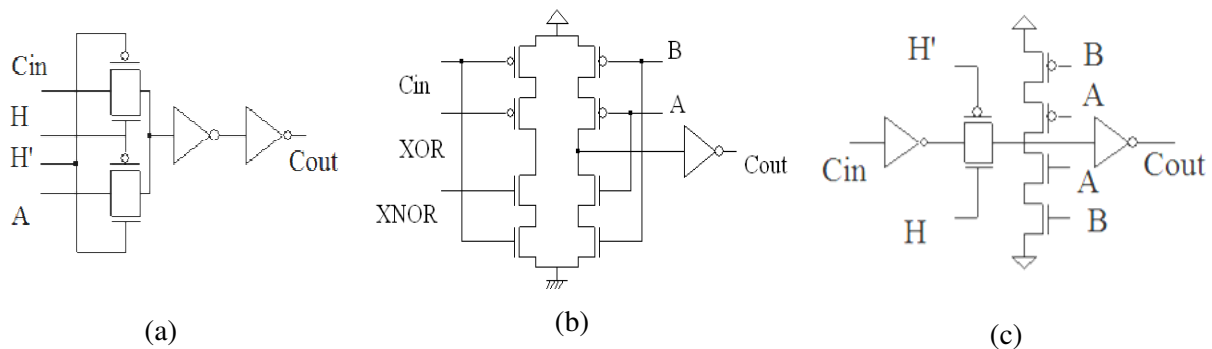


Figure 5: (a) multiplexer using transmission gate (b) Multiplexer based on the static-CMOS logic style (c) Multiplexer based on Hybrid-CMOS logic style

Another possibility is to use the complement of the expression, i.e.,

$$\overline{C_{out}} = \overline{A} \cdot H' + \overline{C_{in}} \cdot H$$

In this case, two inverters will be required to invert the A and C_{in} inputs and one inverter at the output. This will result in unbalanced SUM and C_{out} output switching times and extra delay.

A circuit based on the static-CMOS logic style is presented in [8] [22]. This circuit overcomes the problems of the TG multiplexer design. It uses ten transistors and is shown in Fig. 5 (b). This circuit possesses all the features of static CMOS logic style such as robustness to voltage scaling and good noise margins.

We propose a hybrid design for Module-III. We use the inherently low power consuming TG logic style and the robust static-CMOS logic style to create a new hybrid-CMOS circuit. The proposed circuit is shown in Fig. 5 (c). The new circuit also utilizes ten transistors and possesses the properties of both static-CMOS and TG logic styles. The carry is evaluated using the following logic expression:

$$C_{out} = (A \oplus B) \overline{C_{in}} + \overline{A} \cdot \overline{B}$$

A transmission gate preceded by a static inverter is used to implement $(A \oplus B) \overline{C_{in}}$. H and H' are the complementary gate signals to this TG. When H is at logic 1 and H' is at logic 0, this unit propagates the $\overline{C_{in}}$ signal to the output. Two PMOS pull-up transistors in series with two NMOS pull-down transistors are used to generate $\overline{A} \cdot \overline{B}$. Complementary A and B signal are not required. When A and B are at logic 0 they switch ON both PMOS transistor to generate C_{out} and assign in logic 1. When A and B are at logic 1 they switch ON both NMOS transistors to generate $\overline{C_{out}}$ and assign logic 0. At all other times, this section remains OFF. The static inverter at the output produces the desired C_{out} output. Table-II shows the results of proposed circuit when compared to the circuit in [15].

Table 2: Simulation results for the proposed Module-III at 50-MHz Frequency and 2V_{DD}

	Static-CMOS Multiplexer	Hybrid-CMOS Multiplexer
No. of Transistor	10	10
Power (μW)	1.337	1.437
Delay (ns)	0.1829	0.1224

Due to the additional inverter in the proposed design, it consumes slightly more power as compared to the circuit in [15]. There is redundant switching at the input since the complement of C_{in} is generated even if it is not propagated to the output. This can be avoided by placing the inverter after the TG but this causes a problem as charge can leak through the closed TG and cause a reversal of voltage level at the output. This tradeoff has to be made but this guarantees excellent signal integrity without any glitches.

IV. PROPOSED FULL ADDERS

As mentioned earlier in Section, the centralized full adders, both XOR and XNOR circuits are present (both in module I) that generate the intermediate signals H and H' . These signals are passed on to module II and III along with the carry from the previous stage and the other inputs A and B to produce and SUM and C_{out} (for both 1st and 2nd category). For the 3rd category, we use proposed circuits from module-I and III and one existing circuit from Module-II. The experiment procedure and the selection of our proposed model were very adaptive and symmetrical. Selecting the best circuits from each of the module we have created three combinations for three categories and compared it with other three combinations using traditional TG 2 to 1 multiplexer. The combinations are compared in terms of number of transistor used in circuits, power consumption and delay. Thus we test our proposed adder's performance and found it really encouraging. The three categorized adders are shown in Figure 7, 8 and 9 respectively.

In Module-I, the proposed XOR–XNOR circuit requires non-complementary inputs which will show perfect output. The analysis of driving capability is the same as XOR structure. By cascading a standard inverter to the XNOR circuit, we will have a driving output, and the signal level at the output end will be perfect in all cases. The same property is presented in the XNOR structure. Module-II is a transmission-function implementation of XNOR function to generate the SUM' followed by an inverter to generate SUM. This provides good driving capability to the circuit. Due to the absence of supply rails there are no short circuit currents. The circuit is free from the problem of threshold loss

amongst all circuits that are used for Module-II [6]. Module -II employs the proposed hybrid-CMOS output stage with a static inverter at the output. This circuit has a lower PDP as compared to the other existing designs. The static inverter provides good driving capabilities as the inputs are decoupled from the output. Due to the low PDP of module II and module III, the new adder is expected to have low power consumption.

V. RESULT AND DISCUSSION

Using our proposed models we created three categorized designs for hybrid-CMOS adder. First circuit based on XOR-XOR based full adder which belongs to first category. Here proposed XOR circuit is used as Module-I, II and proposed 2 to 1 multiplexer is used as Module-III. Figure 7(a) and 7(b) respectively represent the hybrid-CMOS adder (XOR-XOR based full adder) and output C_{out} and Sum together. Second circuit based on XNOR-XNOR based full adder of second category where proposed XNOR circuit used as Module-I, II and proposed multiplexer used as Module-III. Figure 8(a) and 8(b) represent consecutively the hybrid-CMOS adder (XNOR-XNOR based full adder) and outputs of C_{out} and Sum together. The final circuit based on Centralized full adder which belongs to our last category. Proposed XOR-XNOR circuit used as Module-I; Proposed transmission-function implementation of XOR and XNOR used as Module-II and proposed multiplexer used as Module-III. Figure 9(a) and 9(b) respectively represents the hybrid-CMOS adder (Centralized full adder) and output of C_{out} and Sum together.

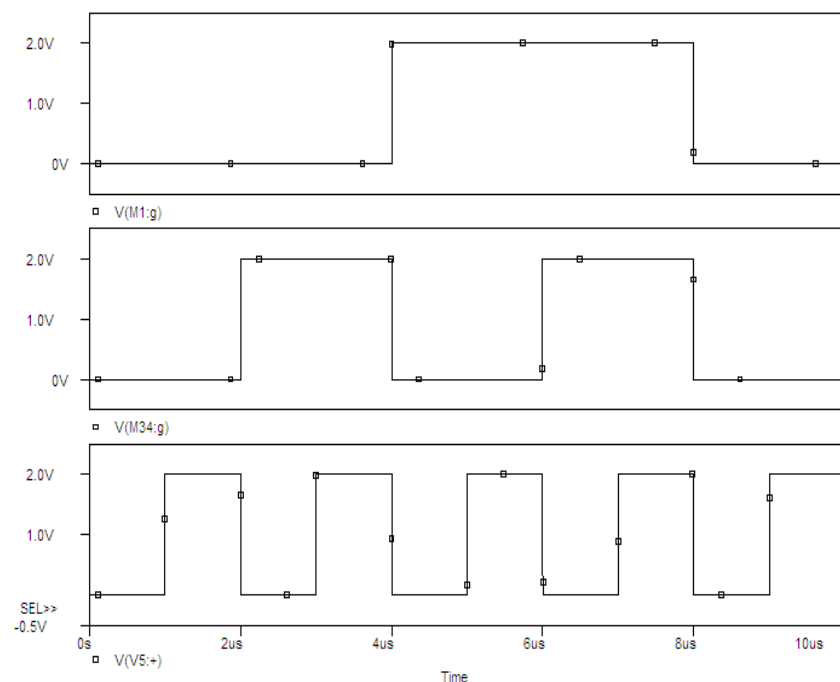


Figure 6: Common input for evaluating all adders

The performance of these three circuits is evaluated based on their transistor numbers, power dissipation and delay. Figure 6 represents the input voltage A , B and C_{in} that used to evaluate all three categorized circuits. Based on our result we finally observed that XOR-XOR based hybrid-CMOS full adder works more efficiently on the basis of all criteria we have mentioned above. Moreover, we have evaluated XOR-XOR based hybrid-CMOS full adder's performances by comparing with all conventional full adders. All simulations are performed using PSPICE, HSPICE and MATLAB.

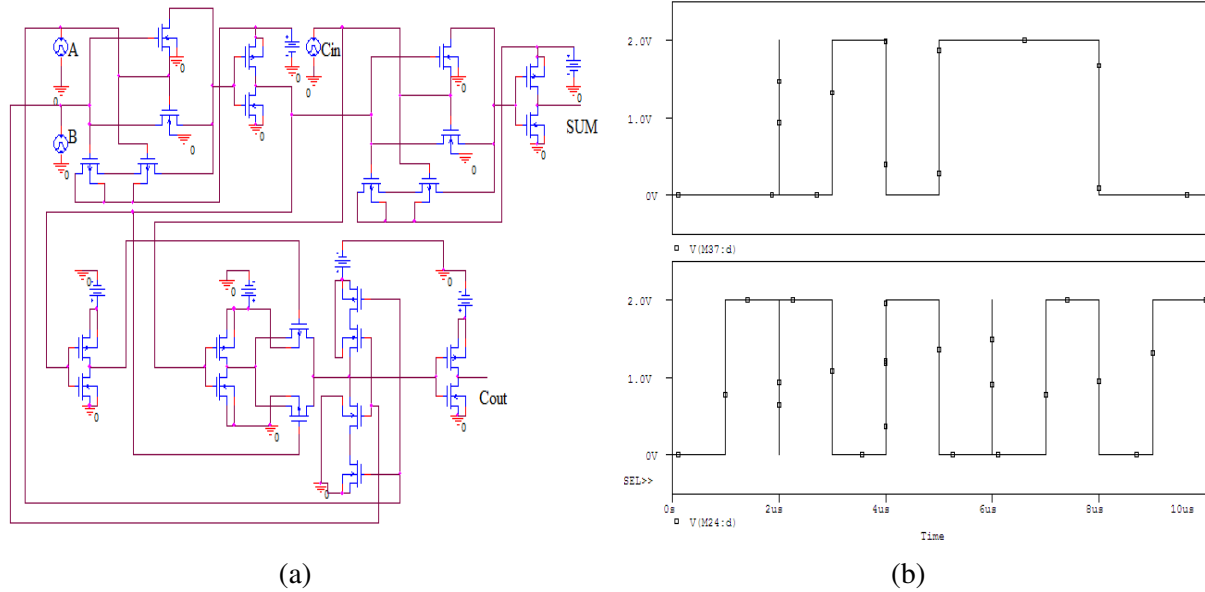


Figure 7: (a) XOR-XOR based Hybrid-CMOS full adder (b) C_{out} and Sum

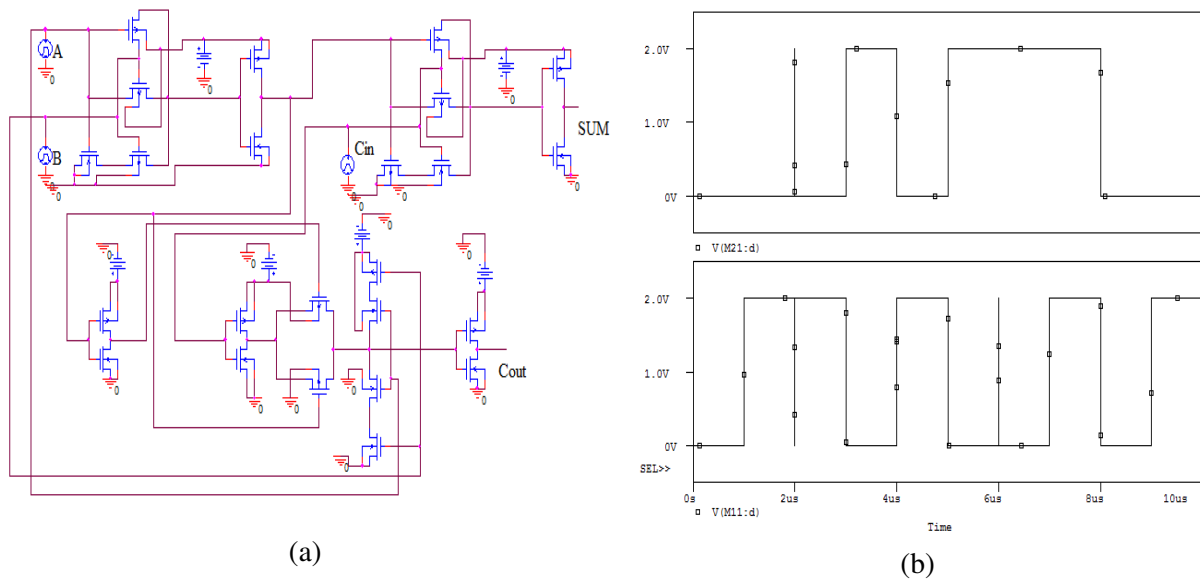


Figure 8: (a) XNOR-XNOR based Hybrid-CMOS full adder (b) C_{out} and Sum

Increase of transistor numbers in chip or digital circuit comes with typical obstacles, even number of transistor may have effect on the overall performance of the circuit. Due to this reason, it was one of our main concerns for designing the full adder without compromising its performance. Three of our proposed designs have twenty four transistors in each and none of them showed any sort of deficiency basis on power dissipation and delay.

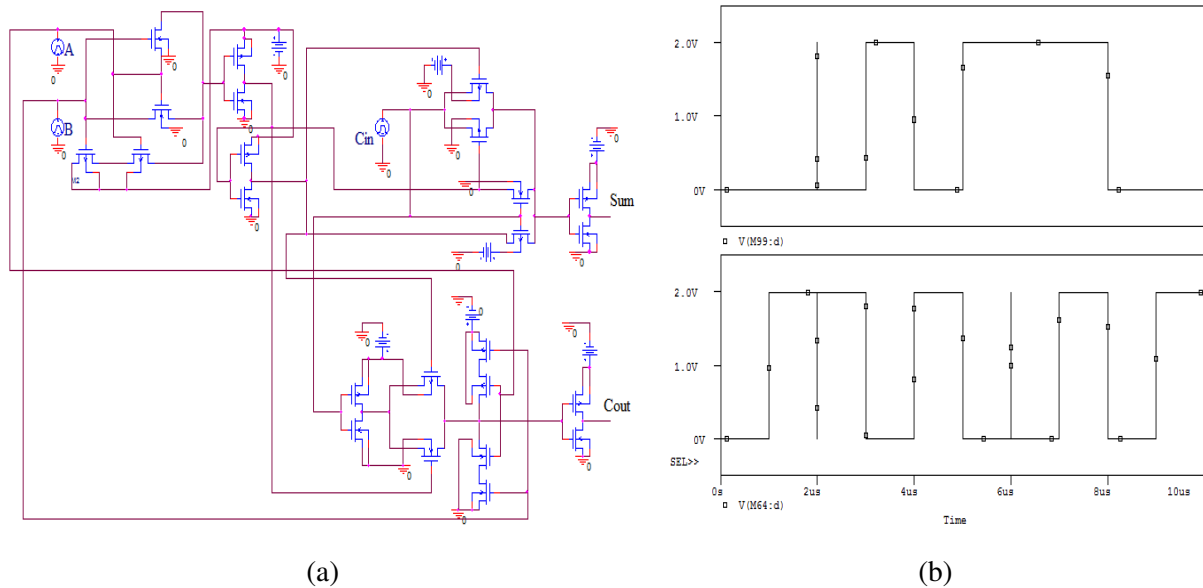


Figure 9: (a) Centralized Hybrid-CMOS full adder (b) C_{out} and Sum

The average power dissipation is evaluated under different supply voltages and different load conditions and is summarized in Figure 10(a) and 10(b) respectively. Among the conventional existing full adders, clearly CPL has the highest power dissipation. The adders TGA and TFA always dissipate less power than others and this can be shown in the graph. Between the two, TGA dissipates lesser power than TFA and the trend continues at low voltages. The degradation in performance of the TFA is higher than the TGA as supply voltage is scaled down. Behind, but closely following the two, comes the static-CMOS full adder. Under varying output load conditions, the adder without driving capability (TGA and TFA) show more degradation as compared to the ones with driving capability (CMOS and CPL). This is as expected since the speed degradation of these designs is highest.

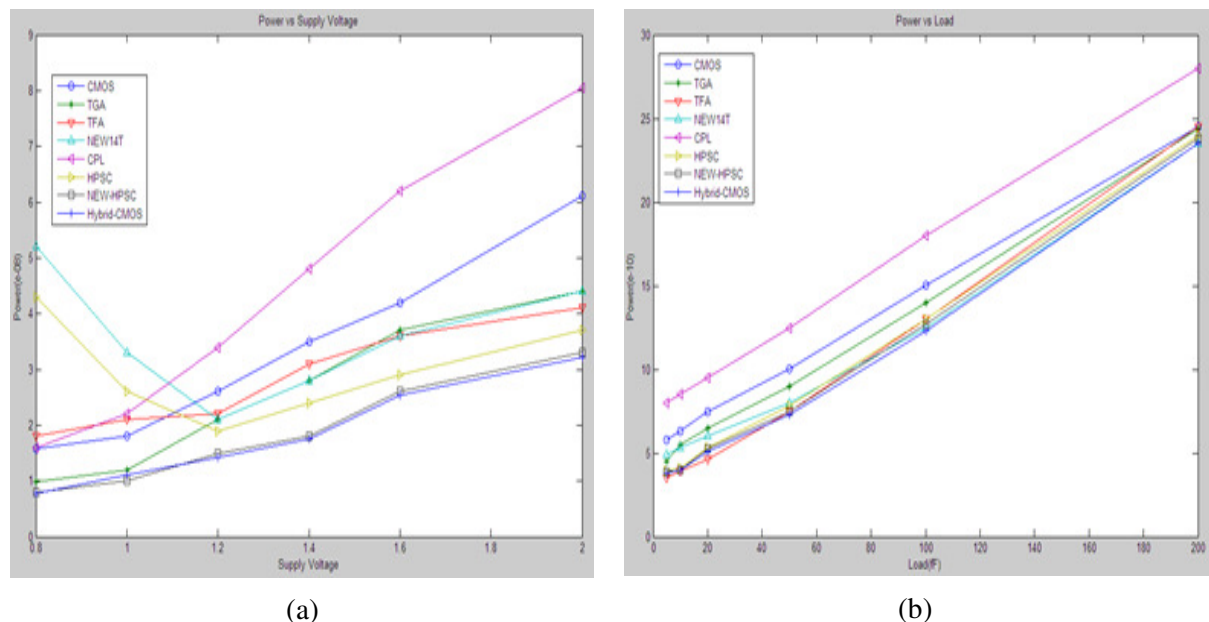


Figure 10: (a) Power vs. Supply Voltage for different full adders (b) Power vs. Load for different full adders

The static-CMOS full adder shows the best performance amongst the conventional full adders under varying load. Among the nonconventional or hybrid-CMOS full adders, the proposed hybrid-CMOS full adder and NEW-HPSC adder have the least power dissipation. The proposed full adder consumes 2% lesser power as compared to the NEW-HPSC adder at $2V_{DD}$ but when the supply voltage is scaled

down, NEW-HPSC adder consumes slightly lesser power. The power dissipation of the proposed adder is roughly 25% less than the next lowest power consuming adder (TGA). With increasing output load, the power dissipation of these adders remains the least as compared to all the considered full adders.

Figure 11(a) and 11(b) respectively represent the delays of full adders at $2V_{DD}$ and load (5.6-200fF). For easy comparison, Table III shows the delay values. From the observation we have learnt that amongst the existing conventional full adders, TGA and TFA (the adders without driving capability) have the smallest delays. TFA has slightly lower delay than TGA at higher supply voltages but the trend reverses at lower supply voltages. The static-CMOS full adder and CPL full adder follow the TGA and TFA adders, CMOS steadily remaining ahead of the CPL adder at each supply voltage. For varying load conditions, TGA and TFA have the low delay at small loads, but the speed degrades significantly at higher loads. Among the existing full adders, CMOS shows the least speed degradation followed by the CPL full adder. This shows that under heavy load conditions, adders with driving capability perform better than those without it (TGA and TFA). Due to these reasons, we compared the proposed hybrid-CMOS adders to the conventional CMOS adders.

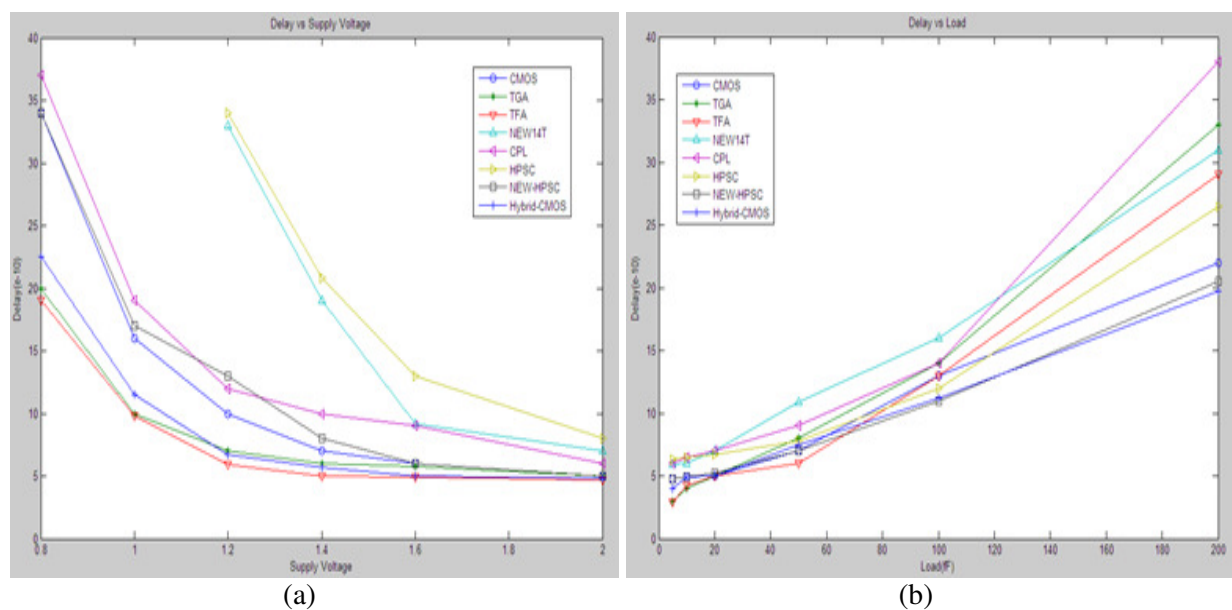


Figure 11:

(a) Delay vs. Supply Voltage for different full adders (b) Delay vs. Load for different full adders

Among the nonconventional or hybrid-CMOS full adders, the proposed hybrid-CMOS full adder shows minimum delay at all supply voltages when compared to the CMOS, HPSC, NEW14T, and NEW-HPSC full adders. At $2V_{DD}$, the proposed adder is 30%, 55%, 88%, and 29% faster than CMOS, HPSC, NEW14T and NEW-HPSC full adders, respectively. At lower supply voltages, the proposed full adder is the fastest. The delay of the proposed hybrid-CMOS adder is slightly higher than TGA and TFA but with increasing load, it displays minimum speed degradation. Overall, when compared to all adders, the proposed adder has minimum speed degradation with varying load.

VI. FUTURE WORK

In recent Years several variants of different logic styles have been proposed to implement 1 bit adder cells [22] [24]. These papers have also investigated different approaches realizing adders using CMOS technology; each has its own pros and cons. By scaling the supply voltage appears to be the most well known means to reduce power consumption. However, lowering supply voltage increases circuit delay and degrades the drivability of cells designed with certain logic style. One of the most important obstacles decreasing supply voltages is the large transistor count and V_{th} loss problem.

In this paper, we used Hybrid CMOS logic style to design our proposed circuit. This type of logic design provides designer flexibility to work on CMOS area to overall performance of a circuit. Different modules give us the opportunity to create new application basis on the requirements. By optimizing the area of CMOS in different modules more efficient designs can be found [19] [23]. But decreasing area size of different modules brings obstacles that can create a negative impact on the overall circuit's performance. So not compromising the negative impact, designer may work on the size and number of the transistors as minimal level as possible. Moreover, a slight improvement in the area of power dissipation, delay, PDP can create huge impact on the overall performance and that can be one of the main concerns for future work. Most of the conventional adders showed lower power consumption at low voltage and higher power consumption at high voltage but our proposed model overcome that obstacle and showed lower power consumption in every kind of input voltage. As different application can be generated using this different modules, designers should take a good look at the power consumption at different input voltage. Another important concern for designing circuits is delay. Decrease of delay and low input voltage might have an impact on the speed of overall circuits. Due to this reason delay is another area where designer can work in future.

VII. CONCLUSION

Hybrid CMOS design style become popular because it provides designer more freedom to work on the performance of single CMOS design to overall circuit. Based upon the application designers can choose required modules as well as efficient circuit from different modules for the implementation. Even by optimizing the transistor sizes of the modules it is possible to reduce the delay of all circuits without significantly increasing the power consumption, and transistor sizes can be set to achieve minimum PDP. Using the adder categorization and hybrid CMOS design style, many full adders can be conceived. As an example, a novel full adder designed using hybrid CMOS design style is presented in this paper that evaluated low power dissipation and delay. The proposed hybrid-CMOS full adder has better performance than most of the conventional full-adder cells owing to the novels design modules proposed in this paper. It performs well with supply voltage scaling and under different load conditions. We recommend the use of hybrid-CMOS design style for the design of high performance circuits.

REFERENCES

- [1]. H.T. Bui, Y. Wang and Y. Jiang, "Design and analysis of low-power 10-transistor full adders using XOR-XNOR gates," IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process, Vol. 49, no. 1, pp. 25-30, Jan. 2002.
- [2]. V. Tiwari, D. Singh, S. Rajogopal, G. Mehta, R. Patel and F. Baez, "Reducing power in high-performance microprocessors" in Proc. Conf. Des. Autom, 1998, pp. 732-737
- [3]. R. Zimmermann and W. Fichtner, "Low power logic styles: CMOS versus pass-transistor logic," IEEE J. Solid-State Circuits, vol. 32, no. 7, p. 1079-1090, July 1997
- [4]. N. Zhuang and H. Wu, "A new design of the CMOS full adder," IEEE J. Solid-State Circuits, vol. 27, no. 5, pp. 840-844, May 1992.
- [5]. N. Weste and K. Eshraghian, "Principles of CMOS VLSI design," in a system perspective, Reading, MA: Addison-Wesley, 1993.
- [6]. A.M. Shams, T. K. Darwish and M.A. Bayoumi, "Performance analysis of low power 1-bit CMOS full adder cells," IEEE Trans. Very Large Scale Integer. (VLSI) Syst, vol. 10, no. 1, pp. 20-29, Feb. 2002.
- [7]. Jyh-Ming Wang, Sung-Chuan Fang and Wu-Shiung Feng, "New Efficient Designs for XOR and XNOR Functions on the Transistor Level", IEEE Journal of Solid-State Circuits, vol. 29, no. 7, July 1994.
- [8]. Summer Goel, Ashok Kumar, Mahdy A. Bayoumi, "Design of robust, energy efficient full adders for deep sub micrometer design using hybrid CMOS logic style,"
- [9]. J. Wang, S. Fang, and W. Feng, "New efficient designs for XOR and XNOR functions on the transistor level," IEEE J. Solid-State Circuits, vol. 29, no. 7, pp. 780-786, Jul. 1994.
- [10]. M. Sayed and W. Badway, "Performance analysis of single-bit full adder cells using 0.18, 0.25 and 0.35 μ m CMOS technologies," in Proc. Int. Symp Circuits Syst. 2002, pp. III-559-III-562.
- [11]. S. Goel, S. Gollamudi, A. Kumar and M. Bayoumi, "On the design of low energy hybrid CMOS 1-bit full-adder cells," in Proc. Midwest Symp, Circuit Syst, 2004, pp. II-209-212

- [12]. H. A. Mahmoud and M. Bayoumi, "A 10-transistor low-power high speed full adder cell," in Proc. Int. ymp. Circuits Syst, 1999, pp. 1-43-46
- [13]. A. Fayed and M. A. Bayoumi, "A low-power 10 transistor full adder cell for embedded architectures," in Proc. IEEE Int. Symp. Circuits Syst, 001, pp. IV-226–229.
- [14]. D. Radhakrishnan, "Low-voltage low-power CMOS full adder," IEE Proc. Circuits Devices Syst., vol. 148, no. 1, pp. 19–24, Feb. 2001.
- [15]. M. Zhang, J. Gu, and C. H. Chang, "A novel hybrid pass logic with static CMOS output drive full-adder cell," in Proc. IEEE Int. Symp. Circuits Syst., May 2003, pp. 317–320.
- [16]. C.-H. Chang, J. Gu, and M. Zhang, "A review of 0.18- μ m full adder performances for tree structured arithmetic circuits," IEEE Trans. Very Large Scale Integer. (VLSI) Syst., vol. 13, no. 6, pp. 686–695, Jun. 2005.
- [17]. H. Lee and G. E. Sobelman, "New low-voltage circuits for XOR and XNOR," IEEE Proc. Southeastcon., pp. 225–229, 1997.
- [18]. H. T. Bui, A. K. Al-Sheraidah, and Y. Wang, "New 4-transistor XOR and XNOR designs," in Proc. 2nd IEEE Asia Pacific Conf. ASICs, 2000, pp. 25–28.
- [19]. Hubert Kaesline, Digital Integrated Circuit Design from VLSI Architectures to CMOS fabrication. Cambridge University Press, New York, 2008
- [20]. S. Goel, M. Elgamel, M. A. Bayoumi, and Y. Hanafy, "Design methodologies for high-performance noise-tolerant XOR–XNOR circuits," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 53, no. 4, pp. 867–878, Apr. 2006.
- [21]. S. Wariya, Himanshu Pandey, R.K.Nagaria and S. Tiwari, "Ultra low voltage high speed 1-bit CMOS adder," IEEE Trans. Very Large Scale Integer, 2010
- [22]. Shiv Shankar Mishra, S.Waria, R.K.Nagaria and S. Tiwari, "New design Methodologies for high speed low power XOR-XNOR Circuits," journal of World Academic Science, Engineering and Technology (WASET) 09, vol. 55, no. 35, pp-200-206, July 2009
- [23]. Digital Design (3rd Edition) by M. Morris Mano, Publisher: Prentice hall: 3 edition, August, 2001
- [24]. R. Pedram, M. Pedram, "Low power design methodologies," kluwer, Norwell, MA, 1996
- [25]. K. Navi, O. Kaehi, M. Rouholamini, A. Sahafi, S. Mehrabi, N. Dadkhahi, "Low power and High performance 1-bit CMOS fill adder for nanometer design," IEEE computer Society Annual Symposium VLSI (ISVLSI), Montpellier fr, 2008, pp. 10-15
- [26]. M. Vesterbacka, "A 14- transistor CMOS full adder with full voltage swing nodes," Proc. IEEE workshop Signal Processing System, October 1999, pp. 713-722

AUTHORS

Mohammad Shamim Imtiaz was born in Dhaka, Bangladesh in 1987. He received his Bachelor degree in Electrical and Electronic Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh in 2009. He is working as a Part-Time Lecturer in the same university from where he has completed his Bachelor degree. Currently he is focusing on getting into M.Sc Program. His research interests include Digital System Analysis and Design, Digital Signal Processing, Digital Communication & Signal processing for data transmission and storage, Wireless Communication.



Md Abdul Aziz Suzon received B.Sc degree in 2011 from Ahsanullah University of Science and Technology, Dhaka, Bangladesh in Electrical and Electronic Engineering .He is working as a Part-Time Lecturer in Ahsanullah University of Science and Technology. Currently he is focusing on getting into M.Sc Program. His research interest includes digital circuit design, VLSI design, Renewable and sustainable energy, Digital communication.



Mahmudur Rahman was born in Dhaka, Bangladesh in 1989. He received his Bachelor degree in Electrical and Electronic Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh in 2011. His research interest includes Digital circuit design, VLSI design, Alternative and renewable energy, Wireless communication, Microcontroller based inverters. Currently he is focusing on getting into Masters Program.



EXAM ONLINE: E-ENABLING EXTENDED LEARNING, ANSWER AND ESSAY EXAMINATIONS

Abdulghader. A. Ahmed, Dalbir S., Ibrahim M.

School of Computer Science, Faculty of Information Science and Technology
University Kebangsaan Malaysia

ABSTRACT

This study reviews the determinant factors for increased motivation of online course. This development in Information and communication technologies (ICT) has leads to major changes in learning-teaching environment. However, teacher's enthusiasm, the roles of instructors warm and friendliness among teaches and students are one of the most important factors for motivation of online course. Students reflections about flexibility are main factors of motivation for online courses independence and freedom of learning can create motivation in an online learning environment. Relevance of course materials, well-planned and organized class sessions, students' active involvement in classroom learning, use of various instructional techniques and illustration with clear examples motivate the students. However, communication and collaboration between students are important factors as they determine the conduciveness of online learning environment/adaptation to technical infrastructure, process of the course and measurement evaluation during online course studies.

KEYWORDS: *E-learning, Exam Online, Motivation, Online community*

I. INTRODUCTION

The integration of Information and communication technologies (ICT) as well as the Internet have contributed immensely to educational changes with flexible, open and more electronically distributed learner-controlled forms of learning (Bossu, Smyth & Stein, 2007). Its widespread and rapid growing significance could transform the educational sectors and influences academic performance. E-learning created new learning/teaching environments system with pedagogical, technological and organizational components focusing on ideal three components to successfully implementation and create balance (Jochems, Merriënboer & Koper, 2004; Garrison and Anderson, 2003). unique strategies to integrate student populations differs online learning across institutions (Hiltz 1993 & Aliva et al. 1997), and national boundaries (Jarvenpaa & Leidner, 1999 and Yoo et al., 2002).

Motivation among student to activate their respective career goal is the main component of the learning environment. Motivation can be as intrinsic and extrinsic however, both form of motivation in learning is very important in students' engagement in the learning experiences. Intrinsic motivation is refers to individual supportive interest, self-requirement, self-determination, self-regulation as well as the autonomy of learning while extrinsic motivation is the external factors that stimulate learners such as behaviours of teachers, learning topics, learning-teaching strategies, teaching-learning process, interaction among students and teachers. Report on motivational perspectives to understand behaviour predict the acceptance of technology. Intrinsic and extrinsic motivation have been found to be key drivers of behavioural intention (Vallerand 1997 & Venkatesh 1999). Woldkowski defined intrinsic motivation as an evocation, an energy called forth by circumstances that connect with what is

culturally significant to the person. Intrinsic motivation is built in learning theories and is used as a constructive measure for user perceptions of technologies (Woldkowski 1993 & Venkatesh 2003). Extrinsic motivation encourages students to commit themselves to instructional goals and however; increases student's achievement earning them reasonable grade or degree. Motivation is a variable that affects student's learning. Students in the virtual learning environment need external motivation in order to stimulate and to support their participation in virtual learning environment. Deci and Ryan (1985) defined extrinsic motivation as the performing of behaviour to achieve a specific reward. From student's perspective, extrinsic motivation on learning may include and not limited to higher grade in exams, awards as well as in prizes winning. Extrinsic motivation could be seen as a factor that influences learning and partly determinant factor to student grade.

Rovai's (2001) reported the need for learning communities and describe four essential elements of classroom community such as spirit, trust, interaction and learning. He stressed that spirit implies the creation of group identity couple with the feeling of belonging to a specific group. Trust he added, is established when group members give honest feedback to others and expect to receive similar feedback. Abundance of research suggests the importance of participant interaction in online learning (Arbaugh 2004; Brower 2003; Shea et al. 2004 & Swan 2003). Mutual interaction exists when students benefit from each members of the group. Students learn when their respective group shares valuable ideas among themselves. However, spirit and trust could pose some definitional and operational challenges such that interaction and learning becomes relatively direct. Participating strategies increases as learning community recognizes the value of interaction and learning online (William Wresch J.B. Arbaugh, & Michael Rebstock 2005). The nature of participant interaction influences and partly determines the level of success in online environments. In contrary, little attention has been paid to examine the nature of interaction across large sample of participants from different online environments. However, this could possibly be as a result of newness of the online learning and the previous online settings.

II. E-LEARNING LEARNING

While building trust, relationships are constrained by the distances that prevent face-to-face meetings and complicated by cultural differences. Kim and Bonk's (2002) studied participation variables among students in Finland, South Korea, and the US and concluded that the range of responses can be seen in students with respect to particular participation practices and culture. The study concludes that Finnish students were more likely to compose group email responses, and more likely to post summaries of comments.

It has been reported that American students participated in email discussions more than their Finnish peers, a result explained by the authors as Finns tend to keep silent and not to speak too much, whereas silence is not habitual with most Americans (Livonen, Parma, Sonnewald & Poole-Kober 1998). Other study asserted that the interactive learning style typical of current classroom conferencing software such as blackboard is most welcomed by peer-oriented learners such as those in the U.S. it was found Asian students relies heavily on direction from their teachers, even in an online environment Liang & McQueen 1993). However, participation rates for Asian students were influenced by faculty involvement, while American students sought regular involvement with respect to their peers. These studies confirm that participation behaviours vary with culture and peers.

Study by Arbaugh et al. (2004) reveals that participation and interaction in distance education formats measures student perceptions of interaction as well as participation. Students can however, underestimate their actual level of participation. Such estimation need not to be the only source of data for participation studies. Online courses could provide archival records of student and instructor participation during course period together with track participation by individuals and groups over the course. Study on the trends by Andrusyszyn et al. (2000) shows those changes in participation rates exist as students grow more accustomed to the technology and task assignments.

III. E-LEARNING COMMUNITY CULTURE

Four essential elements of classroom community were described by Rovai (2000) such as spirit, trust, interaction, and learning. His observations were supported by the importance of trust relationships described by Jarvenpaa et al. (1998), Maznevski et al. (2000) and Leidner (1999).

It has been suggested that online relationships may not be as effective as face-to-face meetings although there are some evidence that personal relationships may develop over time (Chidambaram 1996; Desanctis et al. 1999 & Jarvenpaa 1999). The development of those relationships is constrained further with deadline like end of a course. However, need for efficient communication may take precedence over more relational-based communication. Fundamental aspect of virtual team effectiveness, the presence of personal relationships among the entire team members seems to be more difficult to establish in courses with members that are online.

E-learning provides configurable infrastructure that integrates learning material, tools, and services into a single solution that creates and delivers training or educational content effectively, quickly, and economically (Zhang, Zhou, Briggs, & Nun maker 2006). In many studies, comparisons have been made between the effectiveness of online learning and face-to-face learning. Russell (1999) made an inventory of many of these media comparison studies and concluded that there is no significant difference between the average performances of learners in the case of face-to-face learning compared to learners exposed to distance learning methods. In addition, Ross and Bell (2007) added that this could be dependent on the level of learning found no significant difference in performance at lower levels of abstraction among students in the traditional setting when compared to online students, students in the traditional setting outperformed online students with respect to higher order learning through analysis and synthesizing information.

Internet-based learning provides opportunities for learners to chosen time and location besides; it allows participants to interact with each other with wide range of online resources (Xu & Wang 2006). Based on the nature of materials and interaction with others, online virtual spaces designed for education as well as for training can be either for knowledge construction and group collaboration. Knowledge construction encompasses objectivist and constructivist strategy while collaboration is grouped as individual or group (Benbunan-Fich & Arbaugh 2006). Collaborative activities allow learners greater opportunities for increased social presence and a greater sense of online community with positive online course outcomes (Gunawardena & Zittle 1997).

The combination of knowledge construction with the presence of group collaboration describes four possible web-based learning environments transfer individual, group and constructs individual and group. Besides, anxiety and uncertainty could be reduced as learners communicate with their colleagues (Hiltz et al. 2002). It can be surmised that the participant interaction variables as well as performance depends on the nature of the online environment.

IV. SECURITY IN E-LEARNING

E-learning delivers examinations via a web browser. However, it is important to secure the browser as to prevent student access to the internet, the local file system as well as email. It is important that students entering the E-learning system download and run small windows. This will disable system keys (e.g., ctrl-alt-del, alt-tab, etc.), installs a keyboard hook to trap browser hot-keys which could be used to open new browser windows, launches Internet Explorer in kiosk mode with no address bar, toolbars, or buttons visible or available at the E-learning login page.

After these strategies have been implemented, candidates can navigate and exit the browser by using the interface provided by E-learning. Similar strategy is available using commercial secure browsers such as Respondus LockDown Browser (Respondus 2007). However; once logged, they will be unable to re-login without being provided with additional invigilator password. Therefore, they cannot leave the invigilated environment and re-access the examination.

V. BENEFITS ASSOCIATE WITH ONLINE LEARNING

An effective online learning environment promotes interactivity and collaboration in the learning process. Assessing students' progress in an online environment improves quality and success in Web courses (Hazari et al. 1999). To achieve pedagogical improvements through online learning for teaching and promoting learning, instructors should empower themselves through the use of assessment tools that monitor student's progress (Hazari et al. 1999). The learner-cantered stratgy helps students develop critical thinking skills and allows instructors to assess students' progress (Odin 1997).

Video serves as a sophisticated medium in e-learning because; it is capable of presenting information in an attractive manner. Studies by Wieling (2010) revealed the effectiveness of instructional video on learning outcomes. However, the instructional video used in early studies was primarily either broadcasted through TV programs and on CD-ROM. Recent advances in multimedia and communication technologies have resulted in improved learning systems through the use of video components for instruction.

Carnegie Mellon University just-in-time lecture project observed that video based education and training systems support the same level of teaching and learning effectiveness as face-to-face instruction (Zhang et al., 2006). Online video recordings of lectures allow students to view lectures they have missed or to re-view difficult lectures to improve understanding. Chiu, Lee, and Yang (2006) investigated the viewing behaviour of students in a Chinese grammar course when online post-class lecture videos were made available. They divided students in two groups based on their viewing activity (top 50% and bottom 50%) and found no difference in course grades between the two groups corrected for their GPA.

Additionally, they found that students had a preference for recordings of their own lectures as compared to lectures of a parallel group.

Ross and Bell (2007) on the other compared the performance of students in a quality management course with access to face-to-face lectures as well as the online lecture video recordings to students who only had access to the online lecture recordings. Using a regression analysis they found that the course score of students in the first group with access to the face-to-face lectures was predicted positively by their GPA, negatively by their age, positively by their homework performance and negatively by the number of lectures they viewed online. For students who did not have access to the face-to-face lectures, the course score was positively predicted by their GPA, negatively by their age, positively by their homework performance and positively by the number of lectures they viewed online.

Perceived learning outcome is the observed results in connection with the use of learning tools. Perceived learning outcome was measured with performance improvement, grades benefit; and meeting learning needs. Previous studies shows that perceived learning outcomes and satisfaction are related to changes in the traditional instructor's role in an online learning environment. The recent advances in computer networking technologies and the World Wide Web (Web) break the physical and temporal barriers of access to education. The online learning environment frees students from the constraints of time and place, and it can be made universally available. As online courses improves in educational institutions, assessing students' learning in an online environment is one of the challenges faced by educators.

The Exam Online is currently being improved on the basis of the two live pilots, for future work however, Inclusion of differentiated mark schemes for individual questions, integrated into the marking interface and Offline marking supports personal computers and laptops with later synchronisation however; the main system are helpful. Other useful modifications include the integration with back end system for outputting results. Integration with a free-text computerized marking system provides automatic marking of short answer questions as in Intelligent Assessment Technologies (2007). Support for drawing diagrams when answering questions, potentially on-screen (Thomas 2004) with options for hand written and paper based submission of calculation steps. In addition, simple question and answer measures into the marking process enhances accessibility for sight impaired students areas requiring modification.

VI. LIMITATIONS

The flexibility of asynchronous distance education is valued since students and lecturers need not be online at the same moment however, flexibility is advantageous in an international context where time zones necessarily distribute student's responses. Research examining time intervals for discussion responses could be helpful in this context. Studies by Liang et al. (1999) described cultural differences in participation patterns. To account for the differing cultural differences, the learning experience should develop model of online learning effectiveness based on course software, learning theories, course content, and participant characteristics as well as cultural or institutional characteristics Hiltz & Arbaugh 2003).

Difficulties with establishing trust relationships online as well as variables cultural components of participation behaviours constrain the initiation of international online courses. Online programs provide additional international learning opportunities to their students. Macfayden & Hawkes (2002) tracked six online international education projects and found general satisfaction with the efforts. Troutman (1991) reported that students who feel secure in their own personal use of computers also feel positive toward the use of computers in the schools.

Furst et al. (2004) highlighted that challenges such as personal relationship, adding new members restart the team development process which could disrupt the effort expended by the original team members to develop a team identity and resolve conflicts early in their development. A number of studies of online learning reported that participation patterns in online courses decline as the course progresses (Hiltz & Wellman 1997; Berger 1999; and Arbaugh 2000). Active participation through the program period requires extensive effort. In addition, it was pointed out that increase in the class size makes it more difficult to develop a sense of online community

While most studies conducted at American institutions show strong relationship between learner and instructor, learners interaction and online learning outcomes (Arbaugh 2005), the perceptions and expectations of the German students suggest that the role of participant interaction may not be as strong in German institutions suggesting that a particular need for multi-national studies of the relationship between participant interaction and learning outcomes in online courses (Arbaugh & Hiltz 2003).

Instructors are often challenged with designing online discussion and assignments that encourage students to evaluate information, assimilate information as well as making comparisons and connections (Odin 1997). An assessment tool that monitors student's progress enhances the learning process however; assessment should be a continuous in an online learning environment. I have been asserted that an assessment tool must draw the instructor and students into assessment procedures (Prime 1998). Miller et al. (1998) added that for assessment to be useful as part of a learning process, it must be visible and related to the learning goals with assigned grades or marks for the data collected to measure progress. Educational material and online learning has challenged the effectiveness of the traditional educational approach in universities and other education institutions. Consequentially, these institutions struggle to restructure their strategies in providing education and delivering knowledge. There are great expectations surrounding the development and use of online courses owing to its versatility, flexibility and personalization potential. A strong supportive program office responsible for student advising, faculty support, administrative and financial support, technical support, and orientation of new students however, comprehensive guide is essential for online learning environment Online students should have access to the learning resources available to on-campus students and must also be able to obtain course materials from either their university's online bookstore or from Internet booksellers.

VII. CONCLUSIONS

E-learning electronically support learning and teaching process through computer network that enables transfer of skills and knowledge. E-learning system improves learner's knowledge by providing on-line access to integrated information, advice, and learning experiences. E-learning system has been developed to deliver lectures and summative essay style examinations through appropriate setting. The system supports existing examination processes by providing better and more comprehensive examination experience for an increasingly digital cohort and supports efficient blind marking process. Initial pilots confirmed that the system provides effective and efficient means of deploying traditional essay style examinations on-screen and that it as well improves in many ways upon the existing paper-based process. The system is expected to undergo further development and roll-out as it complexity varies with tradition and cultural.

ACKNOWLEDGEMENTS

The research was funded by The General People's Committee for Higher Education, Bane waleed University, Libya.

REFERENCES

- [1] Andrusyszyn et al., 2000 M. Andrusyszyn A. Moen, C. Iwasiw, T. Ostbye, L.Davie and T. Stovring et al., Evaluationof electronic collaborative international graduate nursing education: The Canada Norway experience, *Journal of Distance Education* 115:. 1–15.
- [2] Arbaugh, 2004 J.B. Arbaugh, Learning to learn online: A study of perceptual changes betweenmultiple online course experiences, *The Internet and Higher Education* 7 (3), pp. 169-181.
- [3] Arbaugh, in press (2005) J.B. Arbaugh, Is there an optimal design for online MBA courses?, *Academy of Management Learning and Education* 4 (2).
- [4] Arkkelin, 2003 Arkkelin, D. (2003). Putting prometheus' feet to the fire: student evaluations of Prometheus in relation to their attitudes towards and experience with computers,computer self-efficacy and preferred learning
- [5] Boomsma, 1987 A. Boomsma, Structural equation modeling by example:Applications in educational, sociological,and behavioural research (pp. 160–188), Cambridge University Press, Cambridge, England.
- [6] Clark, R. C., & Mayer, R. E. (2008). E learning and the science of instruction. San Francisco:Pfeiffer.
- [7] Cox, M. J., & Marshall, G. (2007). Effect of ICT: Do we know what we should know? *Education andInformation Technologies*, 12(2),59–70.
- [8] Daley et al., 2001 B.J. Daley, K. Watkins,S.W. Williams, B. Courtenay, M.Davis and Mike, Exploringlearning in a technology-enhanced environment, *Educational Technology and Society* 4 (3).
- [9] Davies, 2003 R.S. Davies, Learner intent and online courses, *The Journal of Interactive Online Learning* 2 (1).
- [10] Davis et al., 1992 F.D. Davis, R.P. Bagozzi and P.R. Warshaw, Extrinsic and intrinsic motivation to use computers in the workplace ,*Journal of Applied Social Psychology* 22: 1111–1132.
- [11] Deci and Ryan, 1985 E.L. Deci and R.M.Ryan, Intrinsic motivation and self determination in human behavior, Plenum, New York.
- [12] D. Xu and H. Wang, (2006). Intelligent agent supported personalization for virtual learning environments, *Decision Support Systems* 42: 825–843.
- [13] Ding, N. (2009). Computer-Supported Collaborative learning & gender.Ph'd dissertation. University Groningen.
- [14] Diseth, Å. (2001). Validation of a Norwegian version of the approaches and study skillsinventory for students (ASSIST): Application of structural equationmodelling. *Scandinavian Journal of Educational Research*,45(4), 381–394.
- [15] Fink, 2003 L.D. Fink, Creating significant learning experiences: An integrated approach to designing college courses, Jossey-Bass, San Fransisco.
- [16] Furst et al., 2004 S.A. Furst, M. Reeves, B. Rosen and R.S. Blackburn, Managing the life cycle of virtual teams, *Academy of Management Executive* 18: 6–20.
- [17] Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering Education*, 78(7), 674–681.
- [18] Felder, R. M., & Soloman, B. A. (1991). Index of learning styles.
- [19] Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*, 21(1), 103–112.
- [20] Garrison, D. R. & T. Anderson.(2003). E learning in the 21st century: a framework for research andpractice, London: RoutledgeFalmer.
- [21] Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*,77(1), 81.
- [22] Hartnett, J., 1999. Interacting with interactions. Inside Technology Training (July/August); [http://www.ittrain.com/learning/online/7-8-99-learning-nuts bolts.htm](http://www.ittrain.com/learning/online/7-8-99-learning-nuts%20bolts.htm)) Retrieved September 3,1999.
- [23] Hazari, S. and Schnorr, D., 1999. THE Journal 26, p. 11 (June); <http://www.thejournal.com/magaine/current/feat01.html>) Retrieved July 22, 1999.
- [24] Hill, R.B., 1997. The design of an instrument to assess problem solving activities. *Journal of Technology Education* 9, p. 1; <http://borg.lib.vt.edu/ejournals/JTE/jte-v9n1/hill.html>)
- [25] Hiltz, 1993 S.R. Hiltz, Correlates of learning in a virtual classroom, *International Journal ofManMachine Studies* 39:. 71-98.
- [26] Hodgson & Watland, 2004. Hodgson and P. Watland, Researching networked management learning, *Management Learning* 35:. 99 116.
- [27] Irani, 1998 Irani, Communication potential, information richness and attitude: a study of computer mediated communication in the ALN classroom, *ALM Magazine*2
- [28] Jarvenpaa & Leidner, 1999 S.L. Jarvenpaa and A.E. Leidner, Communication and trust in global virtual teams, *Organisation Science*10: 791–815.

- [29] Jarvenpaa Jarvenpaa, K. Knoll and D.E. Leidner, 1998. Is anybody out there? Antecedents of trust in global virtual teams, *Journal of Management Information Systems* 14 (4), pp. 29–64.
- [30] Kim & Bonk, 2002 K.J. Kim and C.J. Bonk, Cross-cultural comparisons of online collaboration, *Journal of Computer Mediated Communication* 8: 1–32.
- [31] Jochems, W.; Merriënboer, J. V & R.Koper. (2004). *Integrated e learning : implications for pedagogy, technology and organization*, New York: Routledge Falmer.
- [32] Kaufman, D. M. (2002). *Teaching and learning in higher education: Current trends*, retrieved from <http://www.sfu.ca/lidc/research/kaufman/LifelongLearning.html>
- [33] MacFayden & Hawkes, MacFayden, L. & Hawkes, B. H. 2002. Report on a survey of current uses of ICTs in Canadian international education activities. Vancouver, BC: University of British Columbia and Canadian Bureau for International Education.
- [34] Miller, A.H., Imrie, B.W. and Cox, K., 1998. *Student assessment in higher education*, Kogan Page, London.
- [35] Neeley, L., Niemi, J.A. and Ehrhard, B.J., 1998. Classes going the distance so people don't have to: Instructional opportunities for adult learners. *THE Journal* 4: 72–73 (November);
- [36] Odin, J.L., 1997. ALN: Pedagogical assumptions, instructional strategies, and software solutions, University of Hawaii at Manoa, Honolulu, HI; http://www.hawaii.edu/aln/aln_te.htm Retrieved September 5.
- [37] R. Benbunan-Fich and J.B. Arbaugh, (2006). Separating the effects of knowledge construction and group collaboration, *Information and Management* 33: 778–793.
- [38] Shea et al., 2004 P.J. Shea, E.E. Fredericksen, A.M. Pickett and W.E. Pelz, Faculty development, student satisfaction, and reported learning in the SUNY learning network. In: T.M. Duffy and J.R. Kirkley, Editors, *Learner-centered theory and practice in distance education: Cases from higher education*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 343–377.
- [39] Swan, 2003 K. Swan, Learning effectiveness: What the research tells us. In: J. Bourne and J.C. Moore, Editors, *Elements of quality online education: Practice and direction*, Sloan Consortium, Needham, MA, pp. 13–45.
- [40] S.R. Hiltz and M. Turoff, 2002. What makes learning networks effective? *Communications of the ACM* 4: 56–59.

AUTHORS INFORMATION

Abdulghader. A. Ahmed completed his undergraduate degree in computer science at 7th October University Bani Walid, Libya in 2001. He is a master candidate in computer science at faculty of computer science & Information Technology University Kebangsaan Malaysia (UKM).



Dalbir Singh received the degree in Computer Science from the Universiti Sains Malaysia, in 2002. He received the Ph.D. degree in Computer Science from the Universiti Malaya in 2009. Currently, he is a senior lecturer at National University of Malaysia. His research interest includes Human Factor in Information System.



Ibrahim Mohamed received the degree in Accounting & Finance from the Liverpool JM University, in 1996. He received the Masters degree in Information Technology from the National University of Malaysia in 1999. Currently, he is a lecturer at National University of Malaysia. His research interests include business data modelling.



NOISE MODELING OF SiGe HBT BASED ON THE CHARACTERIZATION OF EXTRACTED Y- AND Z-PARAMETERS FOR HF APPLICATIONS

Pradeep Kumar and R.K. Chauhan

Department of Electronics & Communication Engineering
M.M.M. Engineering College Gorakhpur-273010, INDIA.

ABSTRACT

In last several decades silicon-germanium (SiGe) technology has come into the global electronics marketplace. Commercial SiGe HBT facilitates transceiver designs and recommends transistor-level performance metrics which are competitive with the best III-V technologies (InP or GaAs), while sustaining strict fabrication compatibility with high yielding, low-cost, Si CMOS foundry processes on large wafers. This work depicts the complete an ample process to model the noise characteristics of a high frequency 0.1 μm SiGe HBT based on a direct parameter extraction technique. A modeling and characterization of noise parameters of Silicon-Germanium Hetrojunction Bipolar transistor is examined in this issue. Initially, Noise in SiGe Hetrojunction Bipolar Transistors is conferred in detail. Later, a linear noisy two-port network and its equivalent circuit model are presented for extracting and characterizing the noise parameters, for example, noise resistance (R_n), optimum source admittance (G_{Sopt} , B_{Sopt}) and minimum noise figure (NF_{min}) along with its modeling significance. In next step, a novel idea that explains the impact of Ge concentration on these noise parameters is also portrayed. The noise characteristics of the SiGe HBTs are advanced to those of III-V semiconductor devices. A corroboration of objective validity of the noise modeling scheme and the extraction noise parameter is accomplished in the form of Y-, and Z-parameters. These results have been validated using a viable numerical device simulator ATLAS from Silvaco International

KEYWORDS: SiGe HBT, R_n , NF_{min} , B_{Sopt} , G_{Sopt} .

I. INTRODUCTION

The multibillion semiconductor industry is rapidly using devices/transistors working in several GHz regions and is pushing to demonstrate useful solid-state transistors, and resultant circuits built from them, capable of operating near the THz regime. There are two major driving forces for SiGe solid-state devices: 1) high frequency communications and radars and 2) various niche THz applications. Recent research has focused on expanding THz options from two-terminal devices (e.g., Schottky diodes) to three-terminal devices (transistors) for both application areas. In high-frequency communications and radars, higher bandwidth transistors are desirable in a number of applications. Optical fiber communications require active amplifiers in decision circuits, multiplexers, and phase-lock loops operating at 100-GHz clock frequency and above. High current-gain and power-gain cutoff frequencies (f_T and f_{max}) are also demanded in microwave, millimeter-wave, and submillimeter wave transceiver designs, where progressive improvements in transistor bandwidth enable the evolution of communications and radars ICs operating to higher frequencies. One of the key concerns in high frequency applications is their noise behavior. Therefore, accurate noise modeling of SiGe HBT is required [1]. SiGe HBTs were first demonstrated in the late 1980s [2]. It quickly became accepted in the field of wireless communication applications, in the form of wireless transceiver circuits because the higher performance than the Si bipolar devices and superior integration level than the III-V devices [3][4][5]. The low noise capability is one of the chief reasons for the success of the SiGe HBT in the field of wireless, RF and optical applications [6][7][8][9].

In past few years, various small-signal SiGe HBT models have been developed using numerous parameter extraction methods with the intention of optimizing their frequency response [10]. Since these SiGe devices reach the cut-off and maximum oscillation frequencies (f_T , f_{max}) beyond 500 GHz (half THz) due to the technology characteristics that's why they are suitable for RF, microwave and optical applications. Over and above, SiGe HBTs are competent devices at low-cost due to their simple coupling with Si technology, in contrast with other technologies (III-V) that offer higher velocities but at higher costs. This is the most important reason why these devices are widely used in electronic industries [9][11].

The defects and non-idealities in semiconductor devices can be computed perceptively by Low-frequency electrical noise. This directly or indirectly impacts the device performance and reliability. Thus, it is of major importance to be able to characterize the noise in semiconductor devices. The interest in low-frequency noise in electronic devices has been motivated by at least two factors. First the theoretical and experimental studies of the noise itself are of major interest. The low-frequency noise has a tremendous impact on devices and circuits. It sets the lower limit of detectable signals, and it converts to phase noise and thereby reduces the achievable spectral purity in communications systems. It is therefore of prime importance to be able to characterize the noise from electronic devices. Equally important is the information the noise carries about the microscopically physical processes taking place. In electronic devices, noise is caused by the random movement of discrete charge carriers, and their interaction with the environment in which they move. Hence, they carry useful information about that environment, e.g., the interior of a resistor or other semiconductor device [12].

Accurate transistor models which describe the high frequency noise behavior of the device are of great importance for the low noise circuit design and moreover, a physics-based equivalent circuit model on the noise behavior of the device. To determine the large number of unknowns of a HBT including the intrinsic elements and the extrinsic capacitances, extraction method based on small-signal π topology is used. Conventional procedures or methods based on simple bias measurements work very well if the extrinsic elements of the HBT have been previously determined. This approach may be used through different procedures—DC, cut-off measurements, or optimization. However, it is often very difficult to accurately determine the values of parasitic elements of the HBT, since the usual DC and cut-off techniques offer poor performance for SiGe HBT devices. In order to avoid this drawback, a new technique has been developed which does not require any additional measurements except for the scattering (S)-parameters at different biases. Linear models with a π topology have been tested to fit the measured S parameters properly. The base resistance, which has a significant impact on the high frequency noise characteristics of the transistor, can be obtained in a consistent way, as an accurate determination of the outer elements simplifies the equivalent circuit to a conventional model [13].

In this paper, an accurate noise model of SiGe HBT is presented by estimating the behavior of its noise parameters. The noise parameters for instance minimum noise figure (NF_{min}) noise resistance (R_n) and optimum source admittance $Y_{s,opt}$ are calculated for this device having 0.1 μm base width. The effect of Ge concentration on these noise parameters is also investigated. Following this motivation, in second section, we discuss various low frequency noise-sources in SiGe HBT. In the next section we introduce a noise model to extract the various noise parameters such as R_n , G_{Sopt} , B_{Sopt} and NF_{min} for analyzing the performance of SiGe HBT in high frequency regime. In the fourth section, we discuss the simulation results based on ATLAS. Finally, in section fifth, we concluded with general observations as well as protrusions of this work.

II. SEMICONDUCTOR LOW-FREQUENCY NOISE SOURCES

2.1 Thermal Noise

Inside ohmic device, the charge carriers at temperature T collide with phonons which in turn cause Brownian random motion with a kinetic energy proportional to T . This yields open circuit voltage fluctuations with zero average value and nonzero rms value. This value is given by [12],

$$v_n = \sqrt{\frac{4hfBR}{e^{\frac{hf}{kT}} - 1}} \quad (1)$$

where v_n is the rms value in Volts, $h = 6.63 \times 10^{-34}$ Js is Planck's constant, $k = 1.38 \times 10^{-23}$ JK⁻¹ is Boltzmann's constant, B is the bandwidth of the system in Hz, f is the center frequency of the band in Hz and R is the resistance in Ohms. Here we consider only the first two terms of a series expansion of the exponential, $e^{\frac{hf}{kT}} - 1 \approx \frac{hf}{kT}$. By using the approximation and converting to voltage spectral density v_n^2/B , we get [12],

$$S_v = 4kTR \quad (2)$$

Hence, the thermal noise is a white noise. In other words, this is a noise with a frequency independent spectrum for frequencies up to the validity of the approximation, $f < kT/h \approx 6250$ GHz at 300 K, or $f \approx 1/(2\pi RC)$, or $f \approx 1/\tau_{\text{coll}} \approx 10^{12}$ Hz. Here C is the parasitic capacitance parallel to R and τ_{coll} the mean time between collisions of free charge carriers. Thermal noise is also identified as Nyquist noise or Johnson noise. Thermal noise is usually the white noise floor studied at high frequencies for the MOSFETs and resistors [12].

2.2 Shot Noise

The corpuscular nature of charge transport causes the shot noise. Walter Schottky discovered shot noise in radio tubes in 1918. He developed what has been recognized as Schottky's theorem. Under steady-state conditions, the time-averaged current is constant, while the arrival times of the electrons are not equally spaced in a tube. This is due to the electrons when they leave the cathode at random times. This leads to fluctuations in the measured current, and, it can be described by simple Poisson statistics. It is mandatory that there is a DC current present or there is no shot noise, and thermal noise would dominate. Shot noise can be observed in for example Schottky-barriers and in PN-junctions. In these places the current results from the random emission of charged particles that are independent and discrete. The short circuit current spectral density is given by [12],

$$S_I = 2qI \quad (3)$$

Where $q = 1.6 \times 10^{-19}$ C and I is the DC-current in Ampere. In PN junctions, the shot noise is white up to a frequency given by the reciprocal of the transit time, i.e., as long as the fluctuations are slower than the rate of recombination. Shot noise is normally the white noise floor. This is observed for the bipolar devices, for example, the HBTs and the lasers [12].

2.3 Generation-Recombination Noise

The fluctuations in the number of free carriers associated with random transitions of charge carriers between energy states cause Generation-Recombination (GR) noise. These random transitions of charge carriers occur mostly between an energy band and a discrete energy level (trap) in the bandgap. For a two terminal sample with resistance R , the spectral densities are depicted as [12],

$$\frac{S_R}{R^2} = \frac{S_V}{V^2} = \frac{S_N}{N^2} = \frac{\langle \Delta N^2 \rangle}{N_0^2} \frac{4\tau_N}{1 + (2\pi f \tau_N)^2} \quad (4)$$

Where, S_V , S_R and S_N are spectral densities of voltage, resistance and number of carriers, respectively. $N_0 = hN_i$ is the average number of free carriers. While τ_N is the trapping time. The resultant is of. The Lorentzian type spectrum is approximately constant below a frequency given by $f = 1/(2\tau_N)$, and rolls off like $1/f^2$ at the higher frequencies. These noise signatures are found in all the device types [12].

III. NOISE MODELING

The analytical expression for R_n , NF_{min} , B_{Sopt} , G_{Sopt} are advantageous for gaining additional intuitive insight into device optimization for noise. This can be accomplished using analytical Y-parameter equations. For this purpose a linear noisy two-port network is demonstrated in figure1 [10]. In order to make such analytical expression practical, the accuracy must be balanced against simplicity of

functional form. The power spectral densities of the input noise current (S_{i_n}), the input noise voltage (S_{v_n}) and their cross-correlation ($S_{i_n v_n}$) are given by [6],

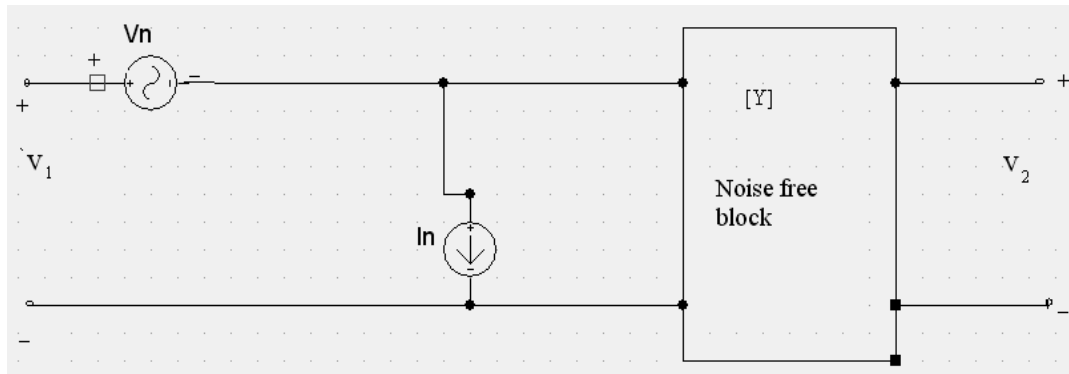


Figure 1: a linear noisy two-port network.

$$S_{v_n} = \frac{V_{na} V_{na}^*}{\Delta f} = \frac{S_{ic}}{|y_{21}|^2} = \frac{2qI_c}{|y_{21}|^2} \quad (5)$$

$$S_{i_n} = \frac{I_{na} I_{na}^*}{\Delta f} = S_{ib} + \frac{S_{ic}}{|H_{21}|^2} = 2qI_B + \frac{2qI_c}{|H_{21}|^2} \quad (6)$$

$$S_{i_n v_n} = \frac{2qI_c y_{11}}{|y_{21}|^2} \quad (7)$$

In further step, we state the Y-parameters in terms of fundamental device parameters, for example β , g_m etc. For the purpose of designing the Niu's method is followed. The small-signal equivalent circuit in simplified is shown in figure 2 [6]. The base resistance is not important for the input impedance at frequencies smaller than f_T . Thus it can be ignored for simplicity, even though it is noteworthy as a noise voltage generator.

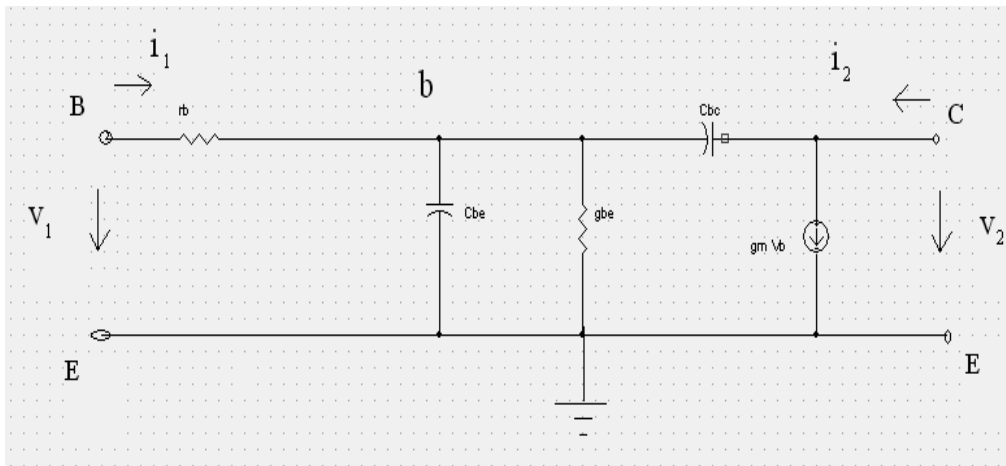


Figure 2: Equivalent circuit for the y-parameter derivation used in analytical noise modeling.

The Y-parameters can be obtained as [6],

$$y_{11} = \frac{g_m}{\beta} + j\omega C_i \quad (8)$$

$$y_{12} = -j\omega C_{bc} \quad (9)$$

$$y_{21} = g_m \quad (10)$$

$$y_{12} = j\omega C_{bc} \quad (11)$$

Where $g_m = qkT/I_C$, and $C_i = C_{be} + C_{bc}$. The C_{be} consists of the EB diffusion capacitance $Si_{1-x}Ge_x$, and $C_{be} = C_{te} + g_m\tau$ with τ being the transit time, and any other EB parasitic capacitances. The C_i is related to f_T and C_{bc} is the total CB junction capacitance, through [6],

$$f_T = \frac{g_m}{2\pi C_i} \quad (12)$$

The oscillation frequency is expressed as [12],

$$f_{\max} = \sqrt{\frac{f_T}{8\pi C_{CB} R_B}} \quad (13)$$

The noise resistance can be determined as [6],

$$R_n = \frac{S_{vn}}{4kT} = r_b + \frac{1}{2g_m} \quad (14)$$

This equation indicates that R_n is directly proportional to the base resistance. R_n also declines with I_C at lower I_C , and then stays constant.

The optimum source admittance can be expressed as [6],

$$G_{s,opt} = \sqrt{\left[\frac{g_m}{2R_n} \frac{1}{\beta} + \frac{(\omega C_i)^2}{2g_m R_n} \left(1 - \frac{1}{2g_m R_n}\right) \right]} \quad (15)$$

$$B_{s,opt} = \frac{-I(S_{i_n v_n^*})}{S_{v_n}} = -\frac{\omega C_i}{2g_m R_n} \quad (16)$$

In general, the admittance increases with collector current and frequency. In the case when diffusion capacitance leads the C_i , then $B_{s,opt}$ becomes independent of I_C , as C_i is proportional to g_m . The absolute value of $B_{s,opt}$ enhances with frequency.

The minimum noise figure is obtained as [6],

$$NF_{\min} = 1 + \frac{1}{\beta} + \sqrt{\left[\frac{2g_m R_n}{\beta} + \frac{2R_n(\omega C_i)^2}{g_m} \left(1 - \frac{1}{2g_m R_n}\right) \right]} \quad (17)$$

$$NF_{\min} = 1 + \frac{1}{\beta} + \sqrt{2g_m r_b} \sqrt{\left[\frac{1}{\beta} + \left(\frac{f}{f_T}\right)^2 \right]} \quad (18)$$

Thus the noise figure NF for two port amplifier with input admittance of Y_s can be given as [14],

$$NF = NF_{\min} + \frac{R_n}{G_s} |Y_s - Y_{s,opt}|^2 \quad (19)$$

Where Y_s is the source admittance and G_s is the real part of Y_s .

IV. SIMULATION RESULTS & DISCUSSION

Based on the above physics based model, the values of various noise parameters are calculated for n-p-n SiGe HBT (figure 3) and investigated for various different Ge concentrations in this paper. Simulation is carried out using ATLAS from SILVACO International. Average Ge concentration in the base region considered in our calculations is varied from 8%-25%. Higher to this are not supported by present epitaxial technologies and beyond it the improvement associated with Ge seizes may be due to lattice constant mismatch [15]. This paper is a next step to our last paper on SiGe HBT [15] for high frequency applications. With the intention of getting excellent accord between analytical and simulated characteristics, all the important physical effects, for example impact ionization (II) is appropriately modeled and accounted for the simulation as well [15].

With the purpose of depicting the complete image of the noise performance, a study of the variation of the noise parameters depending on frequency is employed in figure 4 and 5. The figure 4 describes the dependency and variations of noise parameters (R_n , NF_{\min} , $B_{S,opt}$, $G_{S,opt}$) as a function of frequency. Figure 4(a) shows the variation of minimum noise figure NF_{\min} as a function of frequency and it can be concluded that NF_{\min} of SiGe HBT increases with the increase in frequency. This result matches with the analytical expression of NF_{\min} as in equation (14) which predicts that the NF_{\min} increases monotonically with frequency. It was found that at 65 GHz, simulated NF_{\min} is only 2.70 dB at 65 GHz. This is an admirable result. While at cutoff frequency its value is calculated about 9.82 dB. The variation of optimum source admittance $Y_{S,opt}$ as a function of increasing frequencies are described in figures 4 (b) and (d). The figure 4(b) depicts that behavior of its real part ($G_{S,opt}$) with frequency. It is found that $G_{S,opt}$ increases with the frequency and its calculated value at cut-off frequency is 0.09 mS. Its imaginary part $|B_{S,opt}|$ as a function of frequency is plotted in the figure 4(d) which concluded that the imaginary part is also monotonically increases with frequency. At cut-off frequency, its value is calculated as 0.012 mS. These results are also matched with the analytical expression of $G_{S,opt}$ and $B_{S,opt}$ as in equations (15) and (16). The negative value of imaginary part signifies the requirement of an inductor for reactive noise matching.

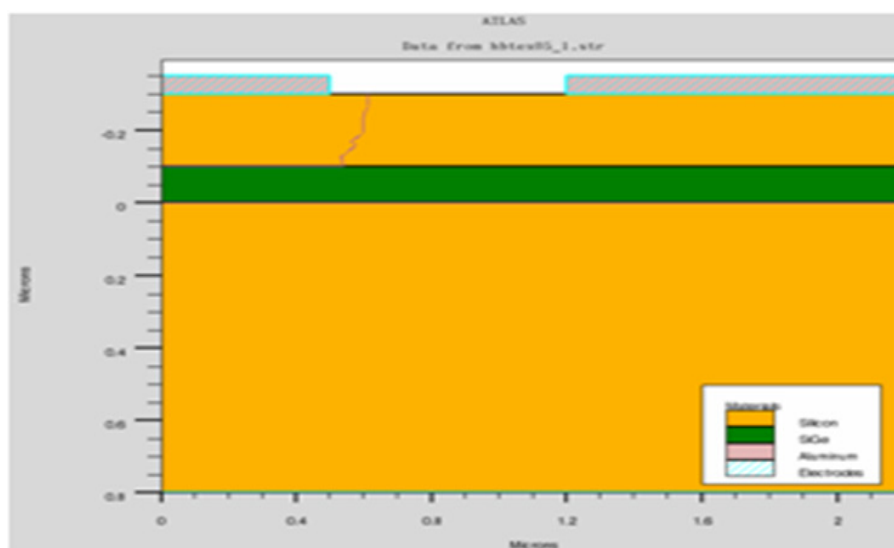


Figure3. The cross-section of the simulated SiGe HBT.

The figure 4(c) demonstrates that the behavior of noise resistance for a frequency range. From the equation (14) it is clear that R_n is directly proportional to base resistance (r_b). From figure 4 (c), it is concluded that noise resistance R_n is weakly depended on frequency. At cut-off frequency, value of R_n is calculated 0.2Ω . This behavior almost fits with its analytical predictions. The following results are demonstrated for maximum oscillation frequency 16.8 THz and corresponding cut-off frequency 13.5 THz [15].

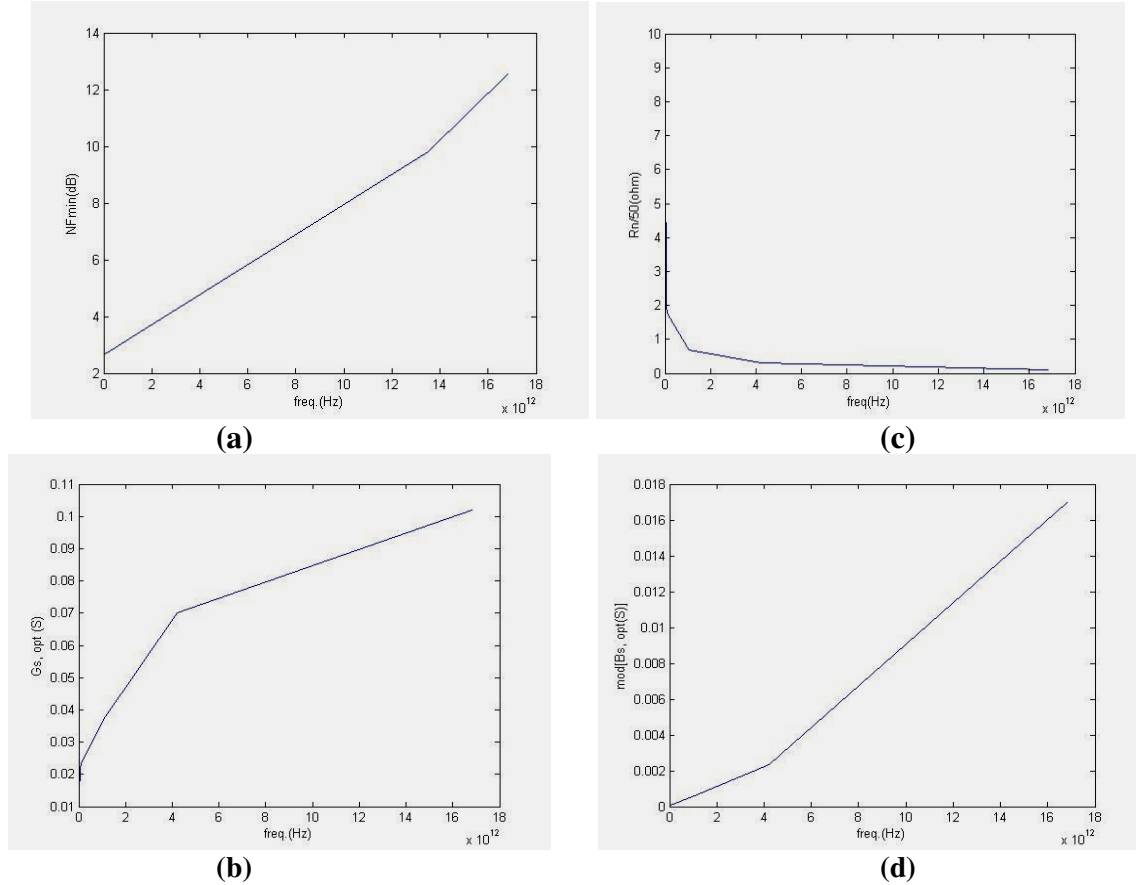
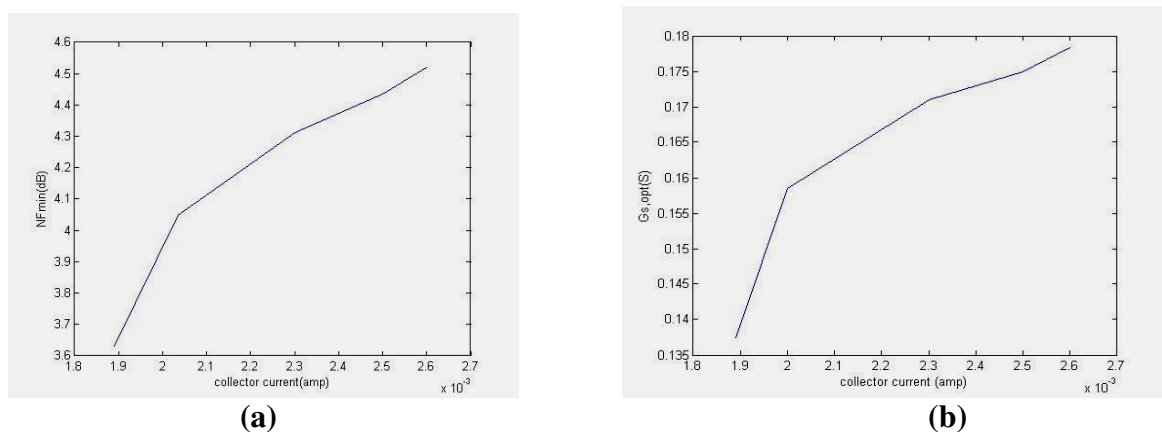


Figure 4. Noise parameters versus frequency for SiGe HBT. (a) NF_{min} vs. Frequency Plot (b) G_{Sopt} vs. frequency plot (c) Noise Resistance vs. frequency plot (d) $mod[B_{Sopt}]$ vs. frequency plot.



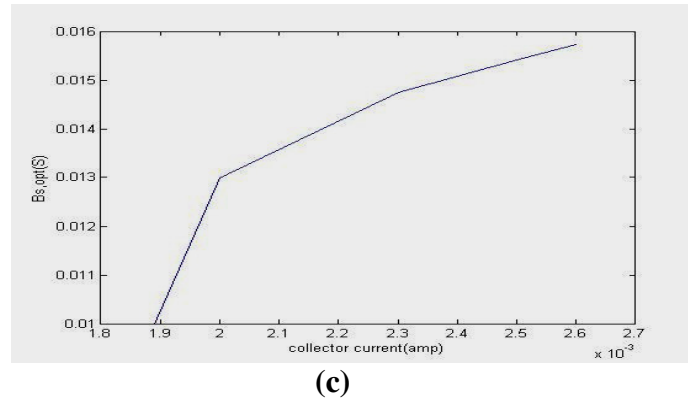


Figure 5. Noise parameters versus collector current (amp.) for SiGe HBT. (a) NF_{min} vs. collector current plot (b) $G_{S,opt}$ vs collector current plot (c) $B_{S,opt}$ vs collector current

The figure 5 exhibits the variations of above noise parameters as a function of collector current. NF_{min} vs. collector current plot is shown in figure 5(a) and it is concluded that the NF_{min} increases monotonically with collector current of SiGe HBT. While the admittance parameters $G_{S,opt}$ and absolute value of $B_{S,opt}$ are also increases with the increment in the collector current of device as shown in figure 5 (b) and 5 (c). These plots of noise parameters vs. frequency and noise parameters vs. collector current are figured with the help extracted Y-, Z- parameters from ATLAS [16].

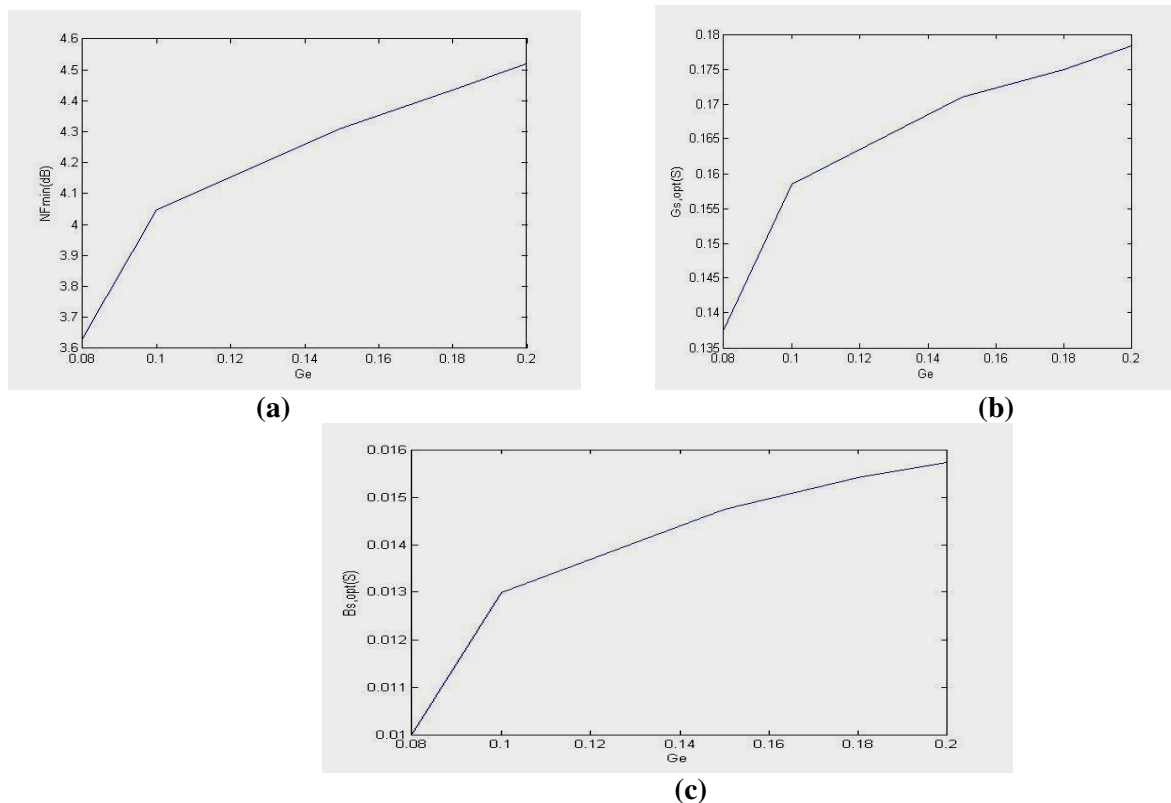


Figure 6. Noise parameters versus Ge concentrations (a) Effect of Ge conc. on NF_{min} (b) Effect of Ge conc. on $G_{S,opt}$ (c) Effect of Ge conc. on $B_{S,opt}$

The effect of germanium concentration over these noise parameters is also investigated in this work. Some observations can be made on the basis of the figure 6 which reveals the impact of Ge concentration on above noise parameters (R_n , NF_{min} , $B_{S,opt}$, $G_{S,opt}$). With this analysis it can be proved the values of noise figure NF_{min} is increased with the increment in Ge concentrations as in figure 6(a). At 0.2 Ge concentration, admirable NF_{min} with fine value of 4.52 dB is achieved. While figure 6(b) and 6(c) display the impact of Ge on optimum source admittance $G_{S,opt}$ and $|B_{S,opt}|$. It is concluded with the help of these two figures that these optimum source admittance parameters are increased with

the increment in Ge concentrations. At cut-off frequency the calculated Noise parameters are summarized in the Table-1 for the noise model of high frequency SiGe HBT.

Table-1: Summary of Noise parameters of SiGe HBT at cut-off frequency

Noise Parameters	Value
$NF_{\min}(\text{dB})$	9.82
$R_n/50 (\Omega)$	0.2
$G_{S,\text{opt}}(\text{mS})$	0.09
Absolute $B_{S,\text{opt}}(\text{mS})$	0.012

Now we will discuss some practical issues of employing this noise model. This model is used for observing the noise behavior of transceiver circuits for mobile wireless communication links because these applications demand highly sensitive circuits and in these applications, the dynamic range and sensitivity of high frequency wireless link depends on HF noise of transistors used in low noise amplifiers [17]. Further, this model can be helpful for estimating the noise performance of millimeter wave- band pass mass market applications, for instance, wireless HDMI and USB, several GHz WLAN and automotive radars [18]. In addition, this proposed model can be useful for approximating the noise characterization that cover the realm of 160 Gb/s fiber optic transmission and MMW imaging [19]. Overall, these noise parameters are extremely valuable for designing the low-signal RF amplifier which results in the high power gain and stable function of the amplifier as well as low noise level in wide frequency range.

V. CONCLUSION

In this work, physics based model and its impact on circuit for low-frequency noise in SiGe HBT has been discussed. In this paper a comprehensive analysis has been done and noise parameters based on equivalent noise model are extracted. It is concluded on the basis of above noise analysis that NF_{\min} increases with frequency. An excellent value of simulated NF_{\min} i.e. 2.70 dB at 65 GHz is achieved. While the Noise Resistance R_n is weakly depend on frequency. On the other hand, $G_{S,\text{opt}}$ and $|B_{S,\text{opt}}|$ increase with frequency and collector current. A novel analysis is also presented which states that the noise figure NF_{\min} as well as the optimum source admittance i.e. $G_{S,\text{opt}}$ and $|B_{S,\text{opt}}|$ of SiGe HBT increases with the Ge contents. At 0.2 Ge concentration, admirable NF_{\min} with fine value of 4.52 dB is attained. This model is used for building the low-signal high frequency amplifier. Such viable noise model can estimate the noise behavior of several GHz WLAN and automotive radars as well as millimeter wave imaging.

REFERENCES

- [1] K. Kumar, and A. Chakravorty, "Physics based modeling of RF noise in SiGe HBTs", IEEE proceedings of International workshop on Electron Devices and Semiconductor Technology IEDST'09', pp. 1-4, 2009.
- [2] G. L. Patton, D. L. Hamee, M. C. J. Stork, B. S. Meyerson, G. J. Scilla and E. Ganin, "SiGe-base, poly-emitter heterojunction bipolar transistors" VLSI Symposium. Technical Digest. pp. 35-36, 1989.
- [3] Han-Yu Chen, Kun-Ming Chen, Guo-Wei Huang and Chun-Yen Chang, "Small-Signal Modeling of SiGe HBTs Using Direct Parameter-Extraction Method", IEEE Transactions on Electron Devices, vol. 53, no. 9, 2006.
- [4] Ankit Kashyap and R.K. Chauhan, "Effect of the Ge profile design on the performance of an n-p-n SiGe HBT-based analog circuit", Microelectronics journal, MEJ: 2554, 2008.
- [5] Pradeep Kumar and R. K. Chauhan, "Electrical parameter characterization of bandgap engineered Silicon Germanium HBT for HF applications", proceedings of International conference on Emerging trends in signal processing and VLSI design, GNEC Hyderabad, Jun. 11-13, pp. 1157-1163, 2010.

- [6] Guofu Niu, "Noise in SiGe HBT RF Technology: Physics, Modeling, and Circuit Implications" Proceedings of IEEE, vol. 93, no. 9, 2005.
- [7] J. D. Cressler, "SiGe HBT technology: a new contender for Si-based RF and microwave circuit applications", IEEE Trans. Microw. Theory Tech., vol. 46, issue 5, 572, 1998.
- [8] Guofu Niu, Zhenrong Jin, John D. Cressler, Rao Rapeta, Alvin J. Joseph, and David Hareme, "Transistor Noise in SiGe HBT RF Technology" IEEE Journal of Solid-state circuits, vol. 36, no. 9, 2001.
- [9] A. Pacheco-Sanchez, M. Enciso-Aguilar and L. Rodriguez-Mendez, "Full comparison between analytical results, electrical modeling and measurements for the noise behavior of a SiGe HBT", IEEE proceedings of ANDESCON-2010, pp. 1 – 5, 2010.
- [10] F. Jun, "Small-signal model parameter extraction for microwave SiGe HBTs based on Y- and Z-parameter characterization", Journal of Semiconductors, vol. 30, no. 8, pp.1-4, 2009.
- [11] N. Zerounian, E. Ramirez-Garcia, F. Aniel, P. Chevallier, B. Geynet and A. Chantre, "SiGe HBT featuring f_T 600 GHz at cryogenic temperature", International SiGe & Ge: materials, processing, and device symposium of the joint international meeting of the 214th meeting of ECS, 2008.
- [12] Jarle Andre Johansen, "Low-frequency Noise Characterization of Silicon-Germanium Resistors and Devices", thesis University of Tromsø, NO-9037 Tromsø, Norway.
- [13] Kenneth H. K. Yau, and Sorin P. Voinigescu, "Modeling and extraction of SiGe HBT noise parameters from measured Y-parameters and accounting for noise correlation", SiRF, p.p.226-229, 2006.
- [14] Neelanjan Sarmah, Klaus Schmalz and Christoph Scheytt, "Validation of a theoretical model for NF_{min} estimation of SiGe HBTs", German Microwave Conference, pp. 265-267, 2010.
- [15] Pradeep Kumar and R. K. Chauhan, "Device Parameter Optimization of Silicon Germanium HBT for THz Applications", International Journal on Electrical Engineering and Informatics, vol. 2, no. 4, pp.343-355, 2010.
- [16] ATLAS User's Manual Device Simulation Software, SILVACO International, 2004.
- [17] M. S. Selim, "Accurate high-frequency noise modeling in SiGe HBTs ", Design Tools/Sftware, pp.24-32, 2006.
- [18] Y. Tagro, D. Gloria, S. Boret, S. Lepillet and G. Dambrine, "SiGe HBT Noise Parameters Extraction using In-Situ Silicon Integrated Tuner in MMW Range 60 - 110GHz", IEEE BCTM 6.1, pp. 83-86, 2008.
- [19] P. Sakalas, J. Herricht, M. Ramonas\$, and M. Schroter, "Noise modeling of advanced technology high speed SiGe HBTs", IEEE proceedings, pp. 169-172, 2010.

Authors Biographies

Pradeep Kumar was born in Allahabad, India in 1985. He received his B.Tech. degree in Electronics & Communication Engineering in 2006. He initially joined VINCENTIT Hyderabad in 2006 and thereafter worked as a lecturer in Dr. K.N.M.I.E.T. Modinagar, Ghaziabad between 2007 and 2008. He is currently pursuing the M.Tech. degree in Digital Systems from Madan Mohan Malviya Engineering College, Gorakhpur, India. His M.Tech. thesis is dedicated towards the modeling and device parameter optimization of Silicon-Germanium HBT for THz applications.



R. K. Chauhan was born in Dehradun, India in 1967. He received the B.Tech. degree in Electronics & Communication Engineering, from G.B.P.U.A.T - Pantnagar, in 1989 and M.E. in Control & Instrumentation, from MNNIT-Allahabad in 1993 and Ph.D in Electronics Engineering, from IT-BHU, Varanasi, INDIA in 2002. He joined the department of ECE, Madan Mohan Malviya Engineering College, Gorakhpur, India as a lecturer, in 1993, as an Assistant Professor since 2002 and thereafter as an Associate Professor since Jan, 2006 to till date in the same institute. He also worked as a Professor in Department of ECE, Faculty of Technology, Addis Ababa University, Ethiopia between 2003 to 2005. He is reviewer of Microelectronics Journal, CSP etc.. His research interests include device modeling and simulation of MOS, CMOS and HBT based circuits. He was selected as one of top 100 Engineers of 2010 by International Biographical Centre Cambridge, England.



DIELECTRIC PROPERTIES OF NORTH INDIAN OCEAN SEAWATER AT 5 GHz

A.S. Joshi¹, S.S. Deshpande², M.L. Kurtadikar³

¹Research Scholar, J.E.S. College, Jalna, Maharashtra, India.

²Rashtramata Indira Gandhi College, Jalna, Maharashtra, India.

³P.G. Department of Physics and Research centre, J.E.S. College, Jalna, Maharashtra, India.

ABSTRACT

This study presents dielectric properties of North Indian Ocean seawater. In all, fourteen seawater samples are collected from Arabian Sea, Lakshadweep Sea, Tip of Bay of Bengal Sea, deep Indian Ocean and Equatorial region. The Von Hippel method is used to measure dielectric properties, both real part ϵ' and imaginary ϵ'' , at 5 GHz and 30 °C using automated C-Band microwave bench set up. The dielectric constant ϵ' and dielectric loss ϵ'' are calculated using least square fitting technique. The salinity measurement of seawater samples are done on autosalinometer. Making use of salinity values of all samples and for 5 GHz and 30 °C, static dielectric constant and dielectric loss are estimated by Klein-Swift model and Ellison et al. model. Experimental and theoretical results are compared. This study emphasizes latitude and longitudinal variations of salinity and dielectric properties. The laboratory data obtained are significant for microwave remote sensing applications in physical oceanography.

KEYWORDS: Seawater Permittivity, Salinity, North Indian Ocean, 5 GHz microwave frequency.

I. INTRODUCTION

Indian Ocean is third largest ocean of the world and has unique geographic setting. The Tropical Indian Ocean (TIO), in particular is significant to oceanographers and meteorologists as it experiences the seasonally reversing monsoon winds and is land locked on northern side. Remote sensing [1-2] of ocean sea surface salinity, sea surface temperature is important in the areas like seawater circulations, climate dynamics, atmosphere modeling, environmental monitoring etc. For microwave remote sensing applications over ocean radar and radiometer, precise values of emissivity and reflectivity are required. The surface emissivity is a complex function of dielectric constant of surface seawater. This complex function is composed of two parts, the real part is known as the dielectric constant (ϵ') and is a measure of the ability of a material to be polarized and store energy. The imaginary part (ϵ'') is a measure of the ability of the material to dissipate stored energy into heat. The two are related by the expression:

$$\epsilon^* = \epsilon' - j\epsilon'' \quad \dots 1$$

The dielectric constant in turn is governed by electrical conductivity and microwave frequency under consideration. The conductivity is governed by salinity and temperature of seawater [3-4]. There are variations in salinity and temperature of ocean resulting variation in dielectric properties and hence in emissivity at that particular location. These variations follow certain pattern latitude and longitude of the location, due to dynamic features of the ocean.

This work focuses on measurement of dielectric properties of seawater samples at 5 GHz at 30°C. The study emphasizes on latitude and longitudinal variations in salinity and dielectric properties. Knowing the dielectric constant and dielectric loss, the parameters like emissivity, brightness

temperature, scattering coefficient can be interpreted, as they are interdependent. Making use of the measured salinity values of all samples, static dielectric constant and dielectric loss are estimated by Klein-Swift model [5-6] and Ellison et al. model [7-8] for 5 GHz and 30 °C. The laboratory data obtained are significant for interpretation of microwave remote sensing applications, and helps in designing active and passive microwave remote sensors.

II. MATERIAL AND METHODOLOGY

2.1. Seawater Sampling

By participating in ORV Sagar Kanya scientific cruise SK-259, organized by NCAOR in May-June 2009 that is summer monsoon period, seawater samples were collected from Arabian Sea, Lakshadweep Sea, Tip of Bay of Bengal Sea, deep Indian Ocean and from equatorial regions of Tropical Indian Ocean. Surface seawater at different locations were drawn through bucket thermometer and two bottles of the samples were preserved around 4°C by standard procedure. Out of two bottles, one of the samples was used to determine the salinity parameter at that location using an Autosalinometer 8400B in the laboratory onboard Sagar Kanya vessel and the other sample of the same location was brought to the Microwave Research Lab, J.E.S. College, Jalna, Maharashtra for dielectric measurement.

2.2. Temperature and Salinity Measurement

The bucket thermometer is used to measure the temperature of surface seawater. Salinity measurements of seawater samples were done using 8400B AUTOSAL onboard ORV Sagar Kanya laboratory. This instrument is semi-portable, semi-automatic and is used in the land based or sea-borne laboratory to determine salinity levels of saline seawater samples and standard seawater sample by measuring their equivalent conductivity. The instrument reading is displayed in terms of conductivity ratio. Inputting the conductivity ratio to the software available in the computer lab, salinity value of the sample is calculated. The software calculates salinity using the following formula. The equation is based on the definitions and the algorithm of practical salinity formulated and adopted by UNESCO/ICES/SCOR/IAPSO Joint Panel on oceanographic tables and standards, Sidney, B.C., Canada, 1980 [9-10].

$$S = \left(a_0 + a_1 R_{15}^{1/2} + a_2 R_{15} + a_3 R_{15}^{3/2} + a_4 R_{15}^2 + a_5 R_{15}^{5/2} + \Delta S \right) \quad \dots 2$$

$$\Delta S = \left[\frac{(T - 15)}{(1 + 0.0162(T - 15))} \right] * \left(b_0 + b_1 R_{15}^{1/2} + b_2 R_{15} + b_3 R_{15}^{3/2} + b_4 R_{15}^2 + b_5 R_{15}^{5/2} \right) \quad \dots 3$$

Where $\sum a_i = 35.0000$, $\sum b_i = 0.0000$,

For, $2 \leq S \leq 42$, and for $-2^\circ\text{C} \leq T \leq 35^\circ\text{C}$.

Table 1. Values of the coefficients a and b

i	a	b
0	0.0080	0.0005
1	-0.1692	-0.0056
2	25.3851	-0.0066
3	14.0941	-0.0375
4	-7.0261	0.0636
5	2.7081	-0.0144

2.3. Measurement of Dielectric Properties

There are several methods of dielectric measurement of liquid [11]. In present work, the dielectric properties of seawater samples are measured using Von Hippel Method [12] for which automated C-

Band microwave bench, as shown in figure 1, is used. The MW bench consists of a low power tunable narrow band VTO-8490 solid-state microwave source; having frequency range of 4.3-5.8 GHz. Tuning voltage is kept at 7 volts, throughout the experiment, which corresponds 5 GHz frequency. The other components of the bench setup are: an isolator, coaxial to waveguide adapter, attenuator, SS tuner, slotted line and the liquid dielectric cell.

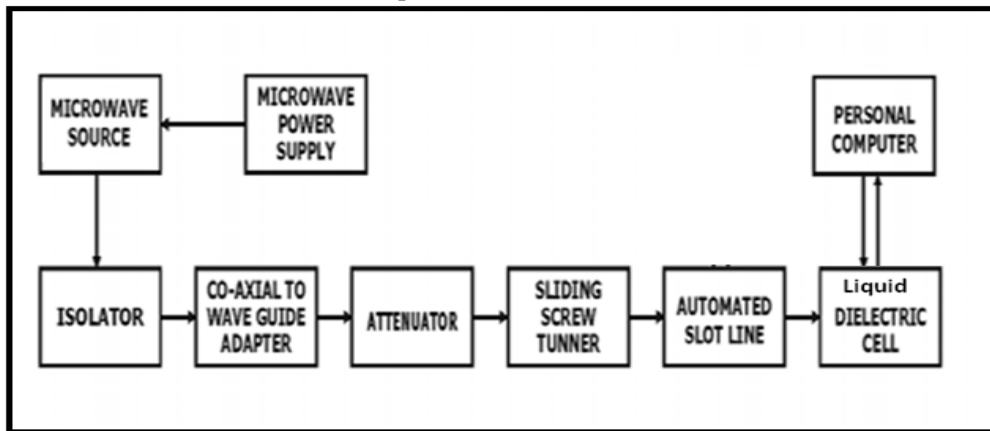


Figure 1. Block diagram of a C-band microwave bench.

Microwave generated by the VTO propagate through the rectangular waveguide to the liquid cell. A desired power level in the line is adjusted with the attenuator. A slotted section with a tunable probe is used to measure the power along the slot line. The crystal detector (1N23) in the probe is connected to a microammeter and to the PC to read, acquire and store the data. The empty liquid dielectric cell is connected at the output end of the bench. The bench is tuned to get symmetrical standing wave pattern in the slot line. The positions of minima are noted from the pattern from which wavelength λ_g of the wave-guide can be calculated. The probe position on the slot line is kept constant at the first minima of the standing wave pattern in the slot line. The liquid dielectric cell is then filled with the sample under consideration. The plunger of the liquid cell is initially set in a position such that the thickness of the liquid column below the plunger is zero. By moving the plunger away from this position, data of microwave power is recorded for different plunger positions. The data of plunger positions and the corresponding power are acquired and stored in a file which is further used to calculate dielectric constant ϵ' and dielectric loss ϵ'' using the least square fit program. The parameters α , β , P_0 , δ are used as the fitting parameters, where α =attenuation factor, β =propagation constant, P_0 =maximum power, and δ =phase factor. The computer program also takes care of calculating error in dielectric constant, $\Delta\epsilon'$, and error in dielectric loss, $\Delta\epsilon''$.

The dielectric properties of seawater samples can be calculated using the relations

$$\epsilon' = \lambda_0^2 \left(\frac{1}{\lambda_c^2} + \frac{(\alpha^2 - \beta^2)}{4\pi^2} \right) \quad \dots 4$$

and

$$\epsilon'' = \frac{\lambda_0^2 \alpha \beta}{2\pi^2} \quad \dots 5$$

where λ_0 is the free space wavelength which can be calculated using the formula

$$\frac{1}{\lambda_0^2} = \frac{1}{\lambda_g^2} + \frac{1}{\lambda_c^2} \quad \dots 6$$

Where $\lambda_c = 2a = 2 * 4.73 \text{ cm} = 9.46 \text{ cm}$, 'a' being the broader side of the C-band rectangular wave-guide.

III. RESULTS AND DISCUSSIONS

The Sea Surface Temperature of collected samples is found to be between 27 °C to 30 °C (Table 1.). Winds over the North Indian Ocean reverse twice during a year. They blow from the southwest during May – September and from the northeast during November – January with the transition-taking place during the months in between. Forced by these winds, circulation in the Indian Ocean has a general eastward direction during summer (May – September) and westward during winter (November – January). During summer, period when seawater samples were collected the monsoon current flows eastward as a continuous current from the western Arabian Sea to the Bay of Bengal [13-14]. These circulations are shown in Figure 2.

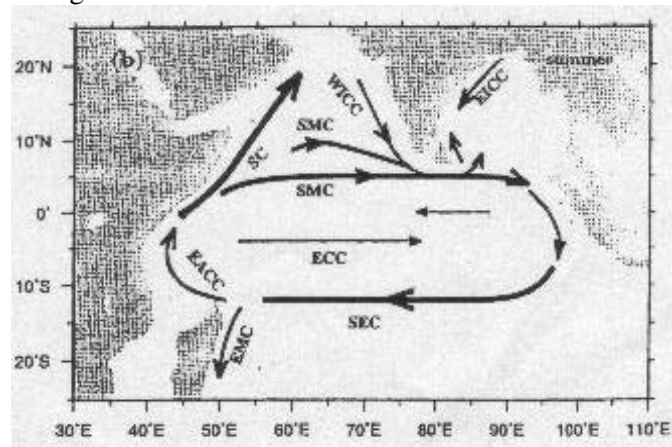


Figure 2. Schematic diagram of major surface currents in the TIO during the southwest (summer) monsoon. The thickness represents the relative magnitude of the current (adapted from Shenoi et al., 1999a) [15].

The Arabian Sea has high salinity (usually in the range 35 to 37) due to excess of evaporation over rainfall. In Table 2, the samples S-01 and S-03 are from Arabian Sea and have higher salinity values compared to other samples.

Table 2. The temperature and salinity values of seawater samples.

Sample	Latitude	Longitude	Temperature °C	Salinity
S-01	N 08° 30'	E 75° 47'	30	35.0238
S-02	N 08° 06'	E 78° 31'	27	34.6434
S-03	N 07° 36'	E 76° 18'	30	35.1564
S-04	N 07° 39'	E 78° 38'	27	34.9782
S-05	N 06° 49'	E 76° 52'	30	34.7117
S-06	N 06° 00'	E 79° 09'	27.5	35.0079
S-07	N 05° 11'	E 78° 00'	29	34.6353
S-08	N 05° 12'	E 79° 39'	28	34.5746
S-09	N 04° 33'	E 78° 25'	29	34.7316
S-10	N 04° 25'	E 80° 16'	28	34.5082
S-11	N 03° 00'	E 81° 20'	27	34.4115
S12	N 02° 46'	E 81° 28'	28.5	34.3808
S13	N 01° 27'	E 82° 24'	28	34.8350
S14	N 00° 37'	E 82° 40'	29	34.9186

In contrast, the Bay of Bengal has much lower salinity due to the large influx of fresh water from river discharge and high amount of rainfall. The samples S-02 and S-04 although located on similar latitude as S-01 and S-03 respectively, differ longitudinal wise and for these Lakshadweep Sea samples, drawn at mouth of Bay of Bengal Sea, decrease in salinity is seen.

The samples S-05, S-07, S-08 are from deep IO and S-06, S-08, S-10, although located on similar latitude, differing in longitude, are towards east, and are from border of Bay of Bengal Sea and Arabian Sea. The salinity values of these samples are found less than the former ones.

As we move towards Equator there is slight decrease in salinity in case of the samples S-11, S-12, but near the equatorial regions, S-13, S-14, a sudden slight increase in salinity value is found. This is due to high evaporation in low-pressure equatorial regions [16].

The dielectric constant ϵ' , dielectric loss ϵ'' , error in dielectric constant $\Delta\epsilon'$ and error in dielectric loss $\Delta\epsilon''$, at 5 GHz, at 30 °C and with varying salinity, latitude and longitude wise in North Indian Ocean are given in Table 3. The magnitude of dielectric constant is found to be 66. It is found that dielectric constant is decreased with increase in salinity. The dielectric loss values are in range of 53 to 58.

Table 3. The experimentally measured values of dielectric constant ϵ' , dielectric loss ϵ'' , error in dielectric constant $\Delta\epsilon'$, error in dielectric loss $\Delta\epsilon''$ of all seawater samples at 5 GHz.

Sample	Latitude	Longitude	Salinity	ϵ'	ϵ''	$\Delta\epsilon'$	$\Delta\epsilon''$
S-01	N 08° 30'	E 75° 47'	35.0238	66.5285	53.8442	6.6474	2.3641
S-02	N 08° 06'	E 78° 31'	34.6434	66.7812	53.9340	7.4309	2.6384
S-03	N 07° 36'	E 76° 18'	35.1564	66.5103	56.9364	6.9828	2.5926
S-04	N 07° 39'	E 78° 38'	34.9782	66.5876	53.8652	7.5291	2.6767
S-05	N 06° 49'	E 76° 52'	34.7117	66.7164	53.9111	6.9207	2.4583
S-06	N 06° 00'	E 79° 09'	35.0079	66.5675	56.4844	7.0274	2.5917
S-07	N 05° 11'	E 78° 00'	34.6353	66.8153	57.7529	6.6314	2.4801
S-08	N 05° 12'	E 79° 39'	34.5746	66.8288	53.9509	7.6745	2.7241
S-09	N 04° 33'	E 78° 25'	34.7316	66.7093	53.9086	6.9227	2.9227
S-10	N 04° 25'	E 80° 16'	34.5082	66.8354	53.9645	7.8615	2.7907
S-11	N 03° 00'	E 81° 20'	34.4115	66.9019	58.549	6.4895	2.4498
S-12	N 02° 46'	E 81° 28'	34.3808	66.954	55.9341	6.0384	2.0061
S-13	N 01° 27'	E 82° 24'	34.8350	66.6813	55.7891	7.2306	2.6381
S-14	N 00° 37'	E 82° 40'	34.9186	66.6072	56.7009	7.5652	2.7969

The values calculated in Tables 4 and 5 are by using Klein and Swift and Ellison et al. models respectively. Comparison of measurement results with these respective models shows that real part, dielectric constant ϵ' values are well in agreement. However, our experimental loss factor is higher by a magnitude of about 20 as compared with the theoretical models. The percentage error in measurement in dielectric constant and loss is of the order of 7 and 2 respectively.

Table 4. The calculated relaxation time, τ (ps), static dielectric constant ϵ_s , dielectric constant ϵ' and dielectric loss ϵ'' using Klein-Swift Model at 5 GHz.

Sample	Latitude	Longitude	Salinity	ϵ_s	τ (ps)	ϵ'	ϵ''
S-01	N 08° 30'	E 75° 47'	35.0238	69.6620	7.0859	66.6042	34.7258
S-02	N 08° 06'	E 78° 31'	34.6434	69.7340	7.0875	66.6715	34.5412
S-03	N 07° 36'	E 76° 18'	35.1564	69.6369	7.0853	66.5807	34.7901

S-04	N 07 ⁰ 39'	E 78 ⁰ 38'	34.9782	69.6707	7.0861	66.6123	34.7037
S-05	N 06 ⁰ 49'	E 76 ⁰ 52'	34.7117	69.7211	7.0872	66.6594	34.5743
S-06	N 06 ⁰ 00'	E 79 ⁰ 09'	35.0079	69.6650	7.0859	66.6068	34.7187
S-07	N 05 ⁰ 11'	E 78 ⁰ 00'	34.6353	69.7355	7.0875	66.6729	34.5372
S-08	N 05 ⁰ 12'	E 79 ⁰ 39'	34.5746	69.7469	7.0877	66.6836	34.5077
S-09	N 04 ⁰ 33'	E 78 ⁰ 25'	34.7316	69.7173	7.0871	66.6559	34.584
S-10	N 04 ⁰ 25'	E 80 ⁰ 16'	34.5082	69.7595	7.0880	66.6953	34.4755
S-11	N 03 ⁰ 00'	E 81 ⁰ 20'	34.4115	69.7777	7.0884	66.7123	34.4284
S12	N 02 ⁰ 46'	E 81 ⁰ 28'	34.3808	69.7835	7.0885	66.7177	34.4135
S13	N 01 ⁰ 27'	E 82 ⁰ 24'	34.8350	69.6977	7.0867	66.6376	34.6342
S14	N 00 ⁰ 37'	E 82 ⁰ 40'	34.9186	69.6819	7.0863	66.6228	34.6747

Table 5. The calculated relaxation time τ (ps) , static dielectric constant ϵ_s , dielectric constant ϵ' and dielectric loss ϵ'' using Ellison et. al. model at 5 GHz.

Sample	Latitude	Longitude	Salinity	ϵ_s	τ (ps)	ϵ'	ϵ''
S-01	N 08 ⁰ 30'	E 75 ⁰ 47'	35.0238	67.9616	7.4153	64.8768	33.8678
S-02	N 08 ⁰ 06'	E 78 ⁰ 31'	34.6434	68.0491	7.4230	64.9537	33.7003
S-03	N 07 ⁰ 36'	E 76 ⁰ 18'	35.1564	67.9311	7.4126	64.8500	33.9261
S-04	N 07 ⁰ 39'	E 78 ⁰ 38'	34.9782	67.9721	7.4162	64.8861	33.8477
S-05	N 06 ⁰ 49'	E 76 ⁰ 52'	34.7117	68.0334	7.4216	64.9399	33.7304
S-06	N 06 ⁰ 00'	E 79 ⁰ 09'	35.0079	67.9652	7.4156	64.8800	33.8608
S-07	N 05 ⁰ 11'	E 78 ⁰ 00'	34.6353	68.0510	7.4232	64.9554	33.6967
S-08	N 05 ⁰ 12'	E 79 ⁰ 39'	34.5746	68.0649	7.4244	64.9676	33.6700
S-09	N 04 ⁰ 33'	E 78 ⁰ 25'	34.7316	68.0288	7.4212	64.9359	33.7391
S-10	N 04 ⁰ 25'	E 80 ⁰ 16'	34.5082	68.0802	7.4258	64.9811	33.6408
S-11	N 03 ⁰ 00'	E 81 ⁰ 20'	34.4115	68.1025	7.4278	65.0006	33.5982
S12	N 02 ⁰ 46'	E 81 ⁰ 28'	34.3808	68.1095	7.4284	65.0068	33.5847
S13	N 01 ⁰ 27'	E 82 ⁰ 24'	34.8350	68.0050	7.4191	64.9150	33.7846
S14	N 00 ⁰ 37'	E 82 ⁰ 40'	34.9186	67.9858	7.4174	64.8981	33.8214

ACKNOWLEDGEMENTS

We are thankful to ISRO for providing the C-Band Microwave Bench Setup under RESPOND project of Dr. M.L. Kurtadikar. Special thanks to NCAOR, Goa, for allowing participation in SK-259 cruise of ORV Sagar Kanya, for seawater sample collection.

REFERENCES

- [1] Fawwaz T Ulaby, Richard K Moore and Adrian K Fung (1986). Vol. 3 Artech House Inc.
- [2] Eugene A. Sharkov (2003). *Passive Microwave Remote Sensing of the Earth*, Springer, Praxis Publishing, UK.
- [3] Smyth, C.P., (1955). *Dielectric Behaviour and structure*, McGRAW-HILL Book company Inc, New York.

- [4] Hasted, J.B, (1973). *Aqueous Dielectrics*, Chapman and Hall Ltd, London.
- [5] Stogryn, A., (1971) Equation for calculating the dielectric constant of saline water, *IEEE transactions on microwave theory and Techniques*, vol.19, pp 733-736.
- [6] Klein, L.A., and Swift, C.T, (1977). An improved model for the dielectric constant of seawater at microwave frequencies, *IEEE J. Oceanic Eng.*, OE-2: pp 104-111.
- [7] Ellison, W., Balana, A., Delbos, G., Lamkaouchi, K., Eymard, L., Guillou, C., and Prigent, C., (1996). Study and measurements of the dielectric properties of sea water, Tech. Rep. 11197/94/NL/CN, European Space Agency.
- [8] Ellison, W., Balana, A., Delbos, G., Lamkaouchi, K., Eymard, L., Guillou, C., and Prigent, C. (1998). New permittivity measurement of seawater,” *Radio Science*, Vol. 33: pp 639-648.
- [9] Lewis, E.L. (1978). Salinity: its Definition and calculation. *J. Geophys. Res.* 83:466.
- [10] Lewis, E.L. (1980). The practical salinity scale 1978 and its antecedents. *IEEEJ. Oceanic Eng.* OE-5:3.pp 14.
- [11] Udo Kaatz (2010). Techniques for measuring the microwave dielectric properties of materials, IOP publishing, *Metrologia*, Vol. 47, pp 91-113.
- [12] Von Hippel A (1954). *Dielectrics & Waves*, Wiley, New York.
- [13] Prasanna Kumar S, Jayu Narvekar, Ajoy Kumar, C Shaji, P Anand, P Sabu, G Rijomon, J Josia, K.A. Jayaraj, A Radhika and K.K. Nair (2004). Intrusion of Bay of Bengal water into Arabian Sea during winter monsoon and associated chemical and biological response, *American Geophysical Research*, vol. 31, L15304, doi:10.1029/2004 GL020247.
- [14] Gangadhara Rao, L.V., Shree Ram, P., (April 2005). Upper Ocean Physical Processes in the Tropical Indian Ocean, monograph prepared under CSIR scientist scheme, National Institute of Oceanography regional centre, Visakhapatnam, pp 4-32.
- [15] Shenoi, S.S.C, Saji, P.K and Almeida, A.M (1999a). Near-surface Circulation and kinetic energy in the tropical Indian Ocean derived from Lagrangian drifters, *J Mar. Res.* Vol. 57, pp. 885-907.
- [16] Shankar .D, Vinayachandra P.N, Unnikrishnan (2002). The Monsoon Currents in the North Indian Ocean, *Progress in Oceanography*, 52(1) pp 63-120.

Authors

Anand Joshi was born in Aurangabad, India in 1981. He received B.Sc. degree in Physics, Mathematics, Computer Science and M.Sc. degree in Physics from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India in 2002 and 2004 respectively. He is currently pursuing a Ph.D.(Physics) degree under the guidance of Dr. M.L.Kurtadikar, Postgraduate Department of Physics and Research Centre, J.E.S. College, Jalna, Maharashtra, India. His research interests include Dielectric measurements, Microwave Remote sensing Applications and Astrophysics.



Santosh Deshpande was born in Parbhani, India in 1974. He received M.Sc. degree in Physics from Swami Ramanand Teerth Marathwada University, Nanded, Maharashtra and M.Phil degree in Physics from Alagappa University, Tamil Nadu, India in 2000 and 2008 respectively. He is currently working as Assistant Professor of Physics in the RMIG College, Jalna, Maharashtra, India. He is also pursuing a Ph.D. degree under the guidance of Dr. M.L.Kurtadikar, Postgraduate Department of Physics and Research Centre, J.E.S. College, Jalna, Maharashtra, India. His research interests include Dielectric measurements, Microwave Remote sensing Applications and Astrophysics.



Mukund L. Kurtadikar was born in Nanded, India in 1951. He received the Master of Science (Physics) and Ph.D. (Physics) degrees from Marathwada University of Aurangabad, India in 1973 and 1983 respectively. He is currently working as Associate Professor of Physics in the Postgraduate Department of Physics of J. E. S. College, Jalna, Maharashtra, India. His research interests include Microwave Remote Sensing Applications, dielectric measurements of soils, seawater, rocks, snow, vegetation etc. He also works on Photometry of Variable Stars using Small Optical Telescope and Scientific Exploration of Historic Monuments. He is a Science Communicator.



AN EFFICIENT DECISION SUPPORT SYSTEM FOR DETECTION OF GLAUCOMA IN FUNDUS IMAGES USING ANFIS

S.Kavitha¹, K.Duraiswamy²

¹Asst.Prof, Nandha Engineering College, Erode, India.

²Dean, K.S.Rangasamy College of Technology, Tiruchengode, India.

ABSTRACT

This paper proposes a computer aided decision support system for an automated detection of glaucoma in monocular fundus images. Identification of Glaucoma using fundus images involves the measurement of the size, shape of the Optic cup and Neuroretinal rim. Optic Cup detection is a challenging task because of the interweavement of cup with the blood vessels. A new color model technique based on pallor in fundus images using K means clustering is proposed to differentiate between the Optic cup to disc boundary. The method differs by initial optic cup region detection followed by the erasure of blood vessels. In addition to the shape based features, textural features are extracted to better characterize the pathological subjects. Optimal set of features selected by Genetic algorithm are fed as input to Adaptive Neuro fuzzy inference system for classification of images into normal, suspect and abnormal categories. The method has been evaluated on 550 images comprising normal and glaucomatous images. The performance of the proposed technique is compared with Neural Network and SVM Classifier in terms of classification accuracy and convergence time. Experimental results shows that the features used are clinically significant for the accurate detection of glaucoma.

KEYWORDS: Optic Cup, Clustering, Glaucoma, Genetic Algorithm, Neuroretinal Rim, ANFIS.

I. INTRODUCTION

Glaucoma the leading cause of blindness and asymptomatic in the early stages and its detection is essential to prevent visual damage [1]. About 2% of the population between 40- 50 years old and 8% over 70 years old have elevated intraocular pressure (IOP) [2] which increases their risk of significant vision loss and blindness. Digital color fundus image has emerged as a preferred imaging modality for large scale eye screening programs due to its non-invasive nature. The less expensive fundus images are used in the proposed work rather than the expensive techniques such as Optical Coherence Tomography (OCT) and Heidelberg Retinal Tomography (HRT).

Optic Disc detection is an important issue in retinal image analysis as it is a significant landmark feature and its diameter is used as a reference for measuring distances and sizes. The optic disc and cup were located by identifying the area with the highest average variation in intensity among adjacent pixels [3]. Automatic detection of optic disc is performed by region of interest based segmentation and modified connected component labelling. Boundary tracing technique was applied to detect the exact contour of optic disc. A quantitative analysis is performed on the neuroretinal rim area to assess glaucoma. [4]. In this approach, a potential set of pixels belonging to cup region is first derived based on the reference color obtained from a manually selected point. Next, an ellipse is fit to this set of pixels to estimate the cup boundary. A variant of this method obtains the cup pixels via thresholding of the green color plane [5]. To handle large inter-image intensity variations that arise due to complex imaging, additional information such as small vessel bends ('kinks') which

anatomically mark the cup boundary have been used in [6]. A Deformable model was presented for the detection of Optic Disc and cup boundaries. The method improves snake model and is robust to edges and ill defined edges [7]. Optic disc is detected using local image information around each point of interest in a multi dimensional feature space. Optic cup is detected by making use of the vessel bends at the cup boundary. Bends in a vessel are detected using a region of support concept and a multistage strategy followed by a local spline fitting to find the desired cup boundary. The method captures OD boundary in a unified manner for both normal and challenging cases without imposing any shape constraint on the segmentation result. Segmentation results shows consistency in handling geometric and photometric variations found across the dataset [8].

A deformable model guided by regional statistics is used to detect the OD boundary. Cup boundary detection scheme is based on Lab color space and the expected cup symmetry. This method uses sector wise information and give rise to fewer false positives and hence better specificity. Error value computed is less for a normal image than for a glaucomatous image [9]. Optic disc and cup are extracted in order to determine cup to disc ratio. Optic disc is extracted using a variational level set method and the detected contour is uneven due to influence of blood vessels. Detection of cup boundary was performed using intensity and threshold level set approach. Thresholding techniques produced better results for both high and low risk retinal images. An ellipse fitting is used to smoothen the boundary [10, 11].

Cup to disc ratio was measured using a vertical profile on the optic disc on the blue channel of the color image to diagnose glaucoma. Sensitivity of 80% and a Specificity of 85% [12] were achieved using vertical CDR measurement for seventy nine images. An algorithm to detect glaucoma using mathematical morphology was developed using fundus images. Developed Neural network system identified glaucoma automatically with a sensitivity of 100% and specificity of 80% [13]. A framework for the detection of glaucoma based on the changes in the optic nerve head using orthogonal decomposition method was used in [14]. The changes in the optic nerve head were quantified using image correspondence measures namely L1 norm, L2 norm, correlation and image Euclidean distance. A novel cup segmentation method based on support vector clustering algorithm [15] is described for the purpose of supporting glaucoma diagnosing in ophthalmology. 30 geometric features were computed on the extracted cup region and the technique achieved 94.5% sensitivity and 97.5% specificity when trained with SVM classifier. 3D images are generally not available at primary care centres due to their high cost. Therefore a solution built around these imaging equipments is not appropriate for large scale screening program. An automated classifier is developed based on adaptive neuro-fuzzy inference system (ANFIS) to differentiate between normal and glaucomatous eyes from the quantitative assessment of summary data reports of the Stratus optical coherence tomography (OCT) images. With stratus OCT parameters as input a good discrimination was achieved between the eyes [16]. A novel method [17] is proposed to detect glaucoma using a combination of texture and higher order spectral features from fundus images. Features extracted have a low p value and are clinically significant. An accuracy of more than 91% is obtained with a random forest classifier combined with z score normalization and feature selection methods.

Most of the works related to Glaucoma detection based on fundus images concentrate only on the Cup to Disc Ratio (CDR). CDR was found to be inconsistent sometimes to detect Glaucoma since the patients may have severe visual loss with a small CDR as in Figure 1 and vice versa. Cup/disc ratio staging system does not account for disc size and that focal narrowing of the neuroretinal rim present between the optic disc and optic cup is not adequately highlighted. So a method has been proposed to accurately detect Glaucoma based on CDR, Neuroretinal rim area to find the rim loss and textural features in order to detect pathological subjects correctly.

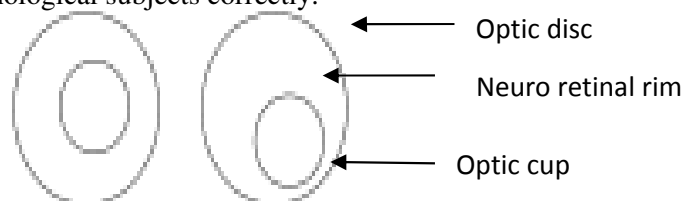


Figure 1. Optic nerve drawings with identical cup/disc ratios but with unequal rim width

Organization of the paper is as follows. Section 3 describes the proposed method for Optic cup detection, feature extraction from Intrapapillary and Peripapillary regions, selection of optimal features by genetic algorithm and the classifier used. Section 4 presents the experimental results and performance analysis of the proposed method. Finally the paper concludes in section 5.

II. MATERIALS USED

Fundus images used in this work are captured by Topcon TRC50 EX mydriatic fundus camera with a 50° field of view at Aravind Eye hospital, Madurai. The image size is 1900x1600 pixels at 24 bits true color image. Doctors in the ophthalmic department of the hospital approved the images for the research purpose.

III. PROPOSED METHOD

An efficient segmentation of optic disc and optic cup is essential to get a better localization of Neuroretinal rim area to diagnose various stages of glaucoma. As glaucoma progresses, the optic cup becomes larger and hence the cup to disc ratio is higher. Further the blood collects along the individual nerve fiber that radiate outwards from the nerve [17]. Such physiological changes are manifested in the fundus images and the experiments shows that the cup to disc ratio and texture features are able to quantify such difference in eye physiology.

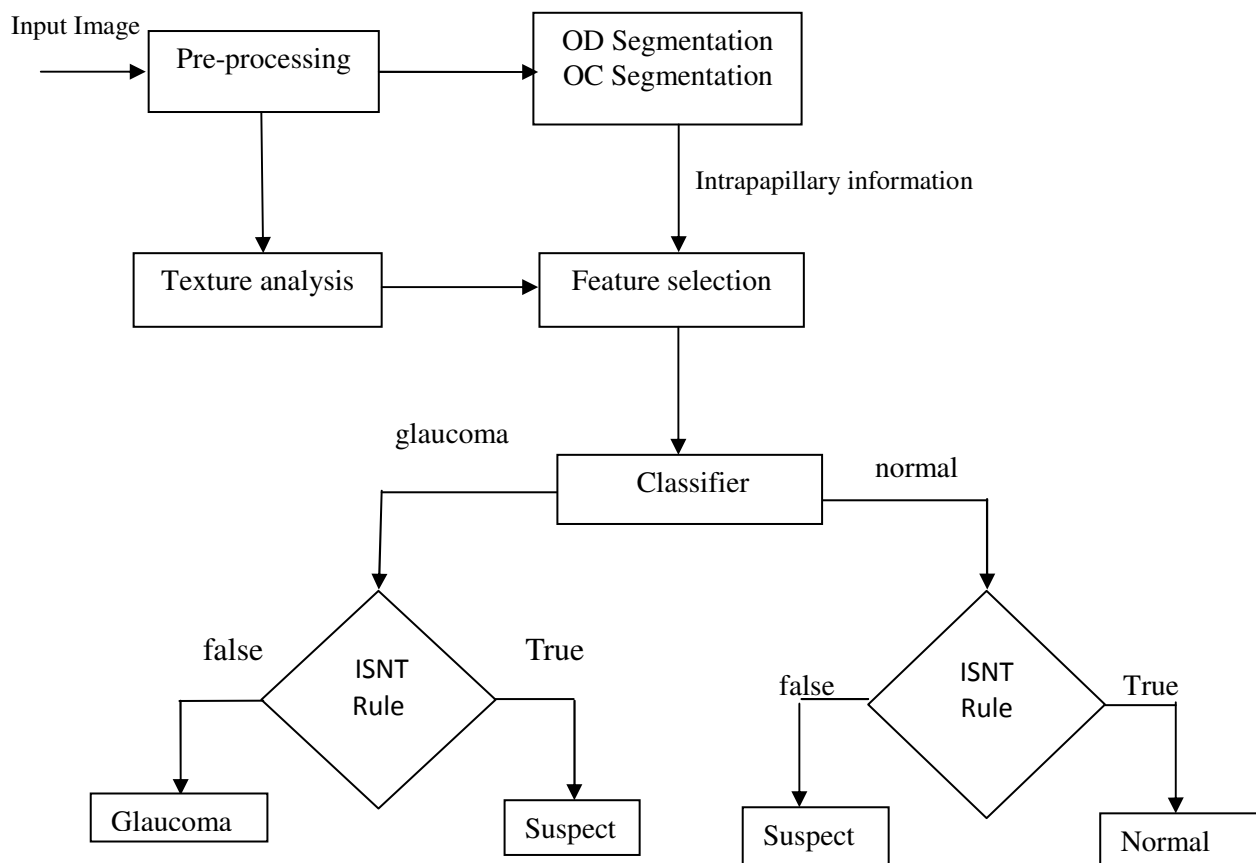


Figure 2. Flowchart of the proposed method

3.1 Pre-processing

The flowchart of the proposed work is shown in Figure 2. RGB color retinal images are pre-processed using anisotropic diffusion filter in order to remove noise. The advantage of anisotropic diffusion [18] is that there is no need to know about the noise pattern or power spectrum previously and also it will

provide better contrast while removing the noise. The filter iteratively uses diffusion equation in combination with information about the edges to preserve edges.

The equation for anisotropic diffusion is defined as:

$$I \operatorname{div}(c(*x, y, t) \nabla I) = c(x, y, t) \nabla I + \nabla c \cdot \nabla I \quad (1)$$

where div is the divergence operator, ∇ is a gradient operator, c is the conduction coefficient function. Anisotropic diffusion filtering introduces a partial edge detection step into the filtering process so as to encourage intra-region smoothing and preserve the inter-region edge. Anisotropic diffusion is a scale space, adaptive technique which iteratively smooths the images as the time t increases. The time t is considered as the scale level and the original image is at the level 0. When the scale increases, the images become more blurred and contain more general information.

3.2 Detection of optic cup

Optic disc is detected using region of interest based segmentation and the bounding rectangle enclosing the region of interest is set as 1.5 times the disc width parameter. In this paper a new approach for the segmentation of Optic cup is proposed. The proposed method shown in Figure 3 is aimed to detect the optic cup exactly to calculate the Neuroretinal rim area present between the disc and cup. Unlike most of the previous methods in the literature, proposed method differs by initial optic cup region detection followed by the erasure of blood vessels to get a higher accuracy.

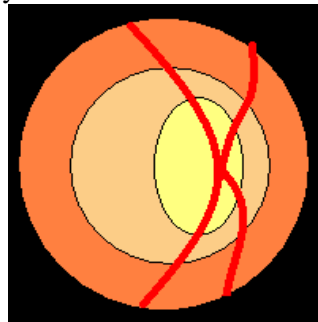


Figure 3. Systematic Representation of the color model

The optic cup and disc areas usually differ in color, known as pallor. This method makes use of this difference in pallor to delineate the cup-disc boundary. Observations on the retinal images show that the actual cup pallor differs between different patients and even between images of the same retina due to changes in the lighting conditions. So the prior knowledge of color intensity of the optic cup cannot be fixed.

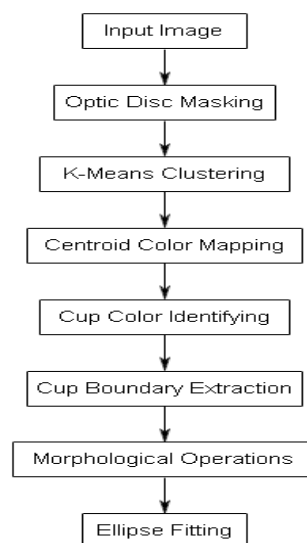


Figure 4. Flow Diagram of the proposed method

Optic cup is detected using the technique proposed in Figure 4. In order to detect the optic cup region from the surrounding region, color space analysis, a segmentation algorithm based on histogram analysis and k means clustering followed by morphological operations has been developed. Since color space transformation plays a significant role in image processing, this step incorporates color information into the segmentation process, where the original RGB image is transformed to different color spaces and it has been found that $L^*a^*b^*$ space consists of a luminosity layer ' L^* ', chromaticity-layer ' a^* ' indicating where color falls along the red-green axis, and chromaticity-layer ' b^* ' indicating where the color falls along the blue-yellow axis. All the color information is in the ' a^* ' and ' b^* ' layers. Optic cup is obtained clearly in this color space when compared with the other color spaces as shown in Figure [5]. These spaces serve as feature vectors for k means clustering. Color difference is measured using Euclidean distance metric.

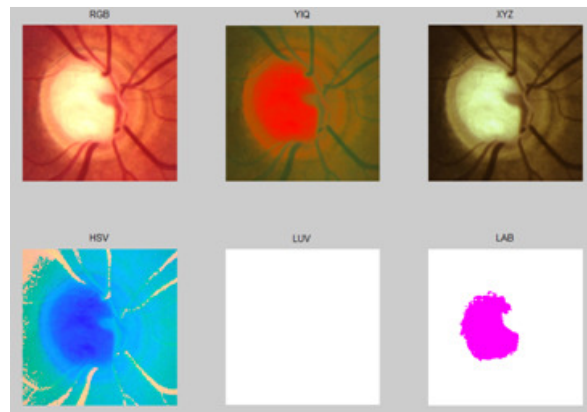


Figure 5. Color space conversion

In the proposed color model for the detection of optic cup, Optic disc consists of regions viz Optic cup, Interior Optic Disc, Exterior Optic Disc and Blood vessels. So the number of clusters is selected as four ($K=4$) manually using domain knowledge. Since the CIE $L^*a^*b^*$ feature space is three dimensional, each bin in the color histogram has N^d-1 neighbors where N is the total number of bins and d is the number of dimensions of the feature space. N is experimented for various values like 5, 10, 15 and the value of N is chosen as 10 by trial and error method and d is 3. Then the 3D colour Histogram is computed. The technique uses histogram information of 3 1D color components to find the number of valid classes. Disc is masked with a radius equal to greater than the size of the optic Disc. The masked image is fed to the clustering process in order to group the pixel values into regions. Number of clusters for k means clustering is determined automatically using Hill Climbing technique[19]. Peaks are identified by comparing the pixels with the neighboring bins and the number of peaks obtained indicates the value of K , and the values of these bins form the initial seeds. The initial seeds for the algorithm was selected from local maximum of the 3D color histogram of the CIE $L^*a^*b^*$ color space. These formed seeds are then passed to K mean clustering. K-Means is an unsupervised clustering algorithm [20] that classifies the input data points into multiple classes based on their inherent distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. K Means Clustering process is explained in the steps below

1. Number of clusters k is taken as four. Lower value of k leads to an increase in the cup size. Higher value results in the predominance of blood vessels. An incorrect value of k gives a sub optimal result.
2. Initialize cluster centres $\mu_1 \dots \mu_k$. Choose k data points and set cluster centres to these points and make them as initial centroids. The data points are grouped into k clusters such that similar items are grouped together in the same cluster
3. For each data point, nearest centroid is found and the data point is assigned to the cluster associated with the nearest centroid. Centroid is the mean of the points in the cluster.
4. Update the centroid of each cluster based on the items in that cluster. The new centroid will be the mean of all points in the cluster.

For a given cluster assignment C of the data points, cluster means m_k is computed as in equation 2.

$$m_k = \frac{\sum_{i:C(i)=k} x_i}{N_k}, k = 1, \dots, K. \quad (2)$$

For a current set of cluster means, each observation is assigned as in equation 3.

$$C(i) = \arg \min_{1 \leq k \leq K} \|x_i - m_k\|^2, i = 1, \dots, N \quad (3)$$

The centroid is taken and data is mapped to the closest one, using the absolute distance between them. 5. The above two steps are iterated until convergence and when there are no new re-assignments it is stopped.

K-means minimizes within-cluster point scatter shown in equation 4

$$W(C) = \frac{1}{2} \sum_{k=1}^K \sum_{C(i)=k} \sum_{C(j)=k} \|x_i - x_j\|^2 = \sum_{k=1}^K N_k \sum_{C(i)=k} \|x_i - m_k\|^2 \quad (4)$$

where x_1, \dots, x_n are data points or vectors or observations

m_k is the mean vector of the k^{th} cluster

N_k is the number of observations in k^{th} cluster

$C(i)$ denotes cluster number for the i^{th} observation.

K-Means clustering groups the pixels within the Optic disc into the above mentioned four regions. Each cluster has a centroid. Then each region is filled with the corresponding region's centroid color. From these 4 regions, the region corresponding to optic cup can be easily identified by its centroid color. Each pixel within a cluster is then replaced by the corresponding cluster centre color. The brightest centroid color corresponds to the optic cup shown in Figure 6. Thus an initial boundary of optic cup is obtained. Pixels that are not classified are assigned to the closest cluster based on a weighted similarity measure between the clusters on the centre and the pixel in the image. $L^* a^* b^*$ color space and k means clustering is more suitable to detect optic cup for normal and pathological subjects and exhibits a high Rand Index and lower Variation of Information (VoI), Global Consistency measure (GCM) and Boundary displacement (BDE) when compared with the other color spaces.

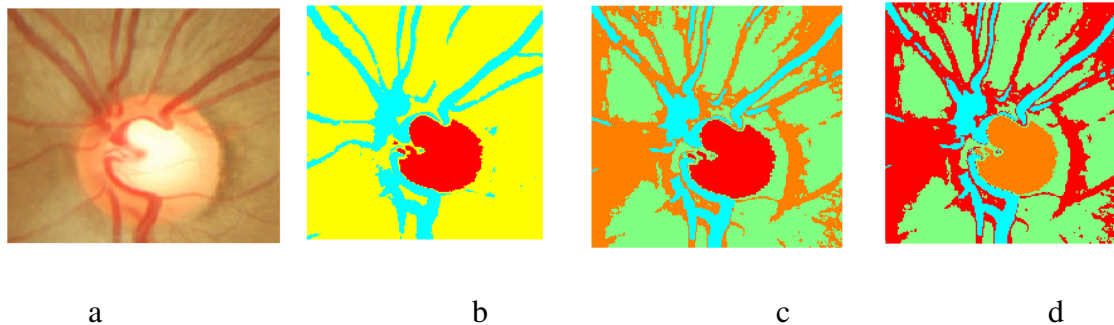


Figure 6. a. Input Image b. clustered outputs for $N=5$ c. $N=10$ d. $N=15$

Boundary of the cup can be obtained using the equation by first eroding the image by the structuring element B and then performing the set difference between A and the eroded image.

$$\beta(A) = A - (A \odot B) \quad (5)$$

The impact of blood vessel region within the cup is removed by morphological operations. This is performed by a dilation followed by erosion operation in the region of interest. A circular window of maximal vessel width as radius is used for dilation and erosion. A 3×3 structuring element is used in this work. Mathematically the functions are expressed using equations 6 and 7.

Dilation of the image A by B

$$A \oplus B = \{ P \in \mathbb{Z}^2 : P = a + b, a \in A \text{ and } b \in B \} \quad (6)$$

Erosion is defined by

$$A \ominus B = \{ P \in \mathbb{Z}^2 : P + b \in A \text{ for every } b \in B \} \quad (7)$$

This step helps to reject outliers inside or outside the cup region and helps to get approximate cup region. Ellipse Fitting Algorithm based on least squares fitting algorithm is used to smooth the cup boundary. The modified optic cup boundary obtained is then fitted with ellipse as in Figure 7. Few sample results for diverse images are shown in Figure 8 for optic cup boundary extraction and Neuroretinal rim area.

3.3 Feature Extraction

Transformation of images into its set of features is known as feature extraction. Features used in this work are based on intrapapillary and peripapillary information from the retinal images. Interpapillary parameters refers to the features extracted from optic disc and optic cup. Cup to disc ratio (CDR) and neuro retinal rim area to disc diameter are extracted from the segmented optic disc and optic cup. CDR was calculated by taking the ratio between the diameter of the Optic cup and disc in the vertical direction. $CDR > 0.3$ indicates glaucoma and $CDR \leq 0.3$ is considered as normal. In glaucoma, structural changes in optic nerve head precede the functional changes. The conventional cup-disc ratio does not measure the actual rim loss which has a more diagnostic value in glaucoma. Neuroretinal rim tissue indirectly indicates the presence and progression of glaucomatous damage and it is related to the disc size. Neuro retinal rim area is calculated by subtracting the area of the optic cup from area of optic disc. Normally the rim is widest in the inferior temporal sector followed by the superior temporal sector, the nasal and the temporal horizontal sector. So Rim to Disc ratio used to estimate the width of the neuroretinal rim is considered as an important feature in the diagnosis of glaucoma. Texture analysis is performed in order to better characterize the abnormal images. Image diagnosis is based on the texture of the segmented portion of the image compared to that of the standard retinal texture image values. Texture extraction is the process of quantifying the texture patterns within a specified neighbourhood of size M by N pixels around a pixel of interest. Features are chosen in order to allow the discrimination between healthy and pathological subjects. The textural properties are derived by using first-order statistics and second-order statistics computed from spatial gray-level co-occurrence matrices (GLCM). GLCM is a second order measure as it includes the relationship between the neighbourhood pixels. For an image of size $m \times n$, a second order statistical textural analysis is performed by constructing GLCM [21]. The data is normalized and contains feature vectors computed around each pixel. The normalized feature vector contains altogether 12 features computed over the window size of ' $n \times n$ ' pixel matrix. Texture analysis is used to estimate the peri papillary information. Normal, suspect and abnormal classes are represented using relevant and significant features to classify the input images. In order to avoid the problem of dimensionality, it is desirable to have a smaller feature set. Twelve features are used in this work among which two are extracted from the segmented optic disc and cup region and 10 are extracted from the texture analysis. Features used are cup to disc ratio, rim to disc ratio, mean, standard deviation, skewness, kurtosis, contrast, correlation, inverse difference moment, variance, energy and entropy.

3.4 Feature selection

Feature selection refers to the problem of dimensionality reduction of data which consists of large number of features initially. The objective is to choose optimal subsets of features from the image. The sequential forward floating selection (SFFS) algorithm [22] and Genetic algorithm was experimented individually to find the best feature set for classification. The algorithm employs a "plus 1, take away r" strategy. Features are added sequentially to an initially empty feature set but, at every iteration features are also removed if that improves performance. In this way "nested" groups of good features can be found.

Genetic algorithm was used to select the most significant features [23] characterizing the shape of the disc and cup region. Since Genetic algorithms are relatively insensitive to noise, they seem to be an excellent choice for the basis of a more robust feature selection strategy to improve the performance

of classification system. In this work, each of the twelve features are represented by a chromosome (string of bits) with 12 genes (bits) corresponding to the number of features. An initial random population of chromosomes is formed to initiate the genetic optimization. A suitable fitness function is estimated for each individual. The fittest individuals are selected and the crossover and the mutation operations are performed to generate the new population. This process continues for a particular number of generations and finally the fittest chromosome is calculated based on the fitness function. Features with a bit value “1” are accepted and the features with the bit value of “0” are rejected.

Feature set selected from Genetic algorithm provides significant six features namely cup to disc ratio, rim to disc ratio, skewness, contrast, correlation and inverse difference moment. In addition to the six features selected by Genetic Algorithm kurtosis is another parameter selected by SFFS algorithm. Each of the features is normalized between 0 to 1 and the weighted features are used for training and testing of instances.

3.5 Adaptive Neuro-Fuzzy Inference System as Classifier (ANFIS)

Adaptive Neuro Fuzzy Inference Systems combines the learning capabilities of neural networks with the approximate reasoning of fuzzy inference algorithms. ANFIS uses a hybrid learning algorithm to identify the membership function parameters of Sugeno type fuzzy inference systems. The aim is to develop ANFIS-based learning models to classify normal and abnormal images from fundus image to detect glaucoma. An adaptive neural network is a network structure consisting of five layers and a number of nodes connected through directional links. The first layer executes a fuzzification process, second layer executes the fuzzy AND of the antecedent part of the fuzzy rules, the third layer normalizes the fuzzy membership functions, the fourth layer executes the consequent part of the fuzzy rules and finally the last layer computes the output of the fuzzy system by summing up the outputs of the fourth layer [24]. Each node is characterized by a node function with fixed or adjustable parameters. Learning or training phase of a Neural network is a process to determine parameter values to sufficiently fit the training data. Based on this observation, a hybrid-learning rule is employed here, which combines the gradient descent and the least-squares method to find a feasible set of antecedent and consequent parameters.

In order to obtain a set of rules and avoid the problems inherent in grid partitioning based clustering techniques, subtractive clustering is applied. This technique is employed since it allowed a scatter input-output space partitioning. The subtractive clustering is one-pass algorithm for estimating the number of clusters and the cluster centres through the training data. Parameters used for clustering are shown in Table 1.

Table 1. Parameters used for clustering

Range of influence	0.5
Squash factor	1.25
Accept ratio	0.5
Reject Ratio	0.15

IV. EXPERIMENTAL RESULTS

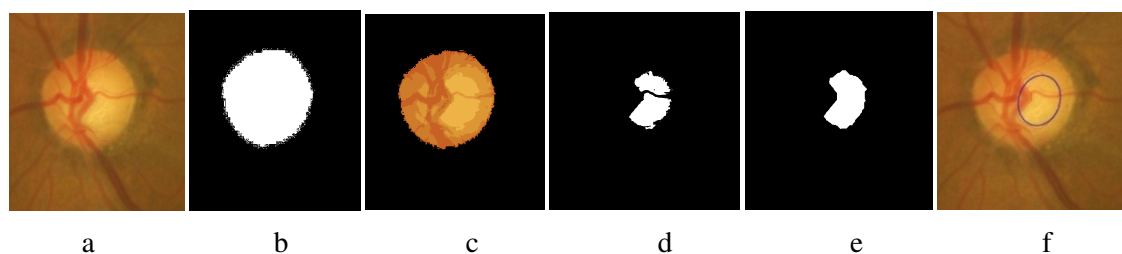


Figure 7. Steps in the detection of optic cup: a. Input image b. Mask image c. Color model d. Initial cup boundary e. Image smoothing f. Ellipse fitting.

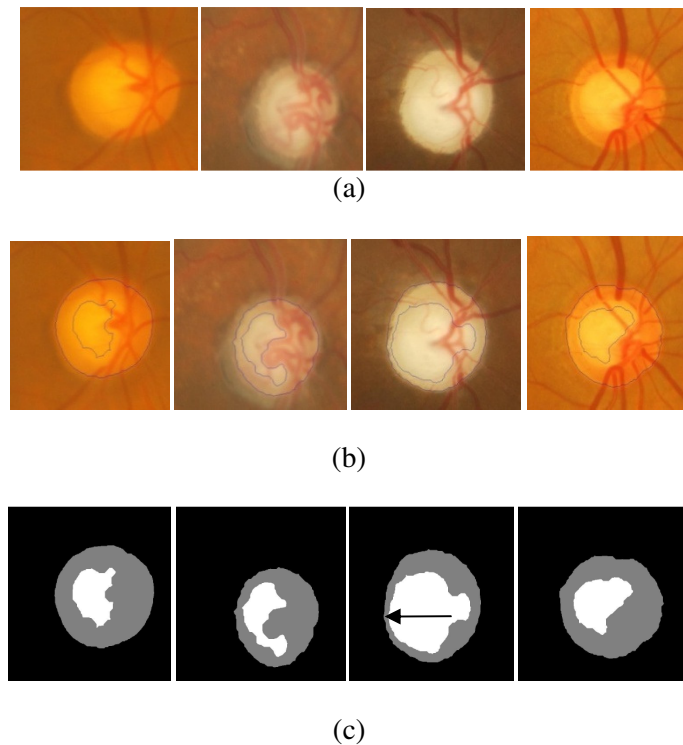


Figure 8. Few sample results a) Input Images b) Cup boundary for the corresponding inputs c) Neuroretinal Rim area present between the disc and cup area shown in arrow mark

V. PERFORMANCE ANALYSIS

A) Optic cup detection

i) To assess the area overlap between the computed region and ground truth of the optic cup pixel wise precision and recall vales are computed

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

where TP is the number of True positives, FP is the number of false positive and FN is the number of false negative pixels.

ii) Another method of evaluating the performance is using F Score given by

$$F = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

Value of F score lies between 0 – 1 and score will be high for an accurate method. Average F score for thresholdind and component analysis are compared and listed in Table 2.

Table 2. F score for cup segmentation

Images	Threshold	Component analysis	Proposed approach
1	0.67	0.72	0.89
2	0.69	0.70	0.86
3	0.66	0.67	0.81

4	0.63	0.73	0.86
5	0.54	0.60	0.78
6	0.71	0.79	0.90
7	0.73	0.78	0.90
8	0.67	0.71	0.86
9	0.68	0.72	0.85
10	0.64	0.76	0.87

B) Performance analysis of the proposed technique

In the proposed system, six features are selected and hence the number of input variables is six. A sample of fuzzy if then rules is framed for fundus images classification. In a fundus image, Fuzzy if then rules form the input for the ANFIS architecture. ANFIS is initialized with 100 iterations and 0.001 step size value for parameter adaptation. Dataset used for fundus image classification is shown in Table 3. A 10 fold cross validation of data is used in the proposed work. From the available dataset, data is split into set1 and testing set. Next, set1 is further divided into training and validation set. Then the classifier is trained using training set and tested on validation set. The process is repeated by selecting various combinations of training and validation set. The classifier which gives best performance is then selected and used to get performance in the testing set.

Table 3.Dataset for Fundus Image Classification

Images	Training Data	Test Data	No of Images/Class
Normal	50	130	180
Suspect	50	120	170
Abnormal	50	150	200
Total	150	400	550

In this work 150 images are used for training and 400 images for testing. 150 images, 50 from each of the class for training and 400 images (130 normal, 120 suspect and 150 abnormal) for testing were used for classification. The schematic of the ANFIS structure obtained for the proposed system is shown in Figure9.

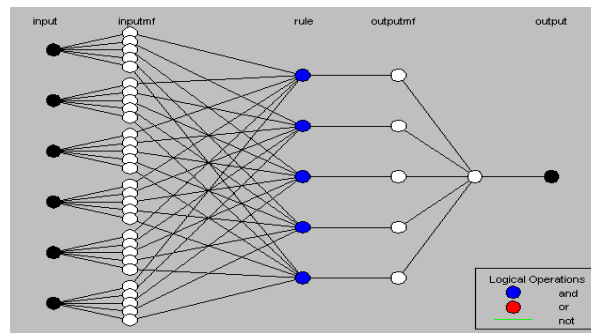


Figure9. ANFIS Structure for the proposed technique

Number of nodes used in the architecture is 79. 35 linear parameters and 60 nonlinear parameters are generated with 5 fuzzy rules. Root Mean square error is 0.022 when testing the data against the FIS structure. Classification accuracy is the ratio of the total number of correctly classified images to the total number of misclassified images. Table 4 shows the performance measure of the classifiers.

Table 4.Classification accuracy results of the classifier

Images	No of test images	ANFIS			Back		
		CCI	MI	CA(%)	CCI	MI	CA(%)
Normal	130	127	3	97.6	125	5	96.1
Suspect	120	119	1	99.1	117	3	97.7
Abnormal	150	149	1	99.3	147	3	98

CCI = Correctly Classified Images, MI = Misclassified Images, CA = Classification Accuracy

Performance of each classifier is measured in terms of sensitivity, specificity, and accuracy. Sensitivity is a measure that determines the probability of results that are true positive such that the person has glaucoma. Specificity is a measure that determines the true negatives that the person is not affected by glaucoma. Accuracy is a measure that determines the results that are accurately classified. The same dataset is used for neural network based Back propagation classifier. MATLAB (version 7.0) is used for implementation of the work. Comparative analysis performed between the classifiers based on correctly classified images, is shown in Table 5. Comparative performance of the classifier using the optimal feature subset selection is shown in Figure 10.

Table 5. Performance measure of the classifiers

Classifier	Specificity(%)	Sensitivity(%)	Accuracy(%)
ANFIS	97.6	99.2	98.7
BACKPROPAGATION	96.1	97.7	97.25

With Performance evaluation classification by means of Area under receiving operating characteristics (ROC), Classification with optimal feature selection achieves 0.99 A_z , 0.0437 standard error and 0.7427 computation seconds for ANFIS and 0.93 A_z with 0.0123 standard error for Back propagation. Classification without optimal feature selection has 0.91 A_z with 0.01405 standard error and 4.16 seconds for computation. Convergence time and RMSE of ANFIS is very less compared to Back Propagation Neural Network. ANFIS gives a good classification performance when compared to back propagation in terms of convergence time and Root mean square error.

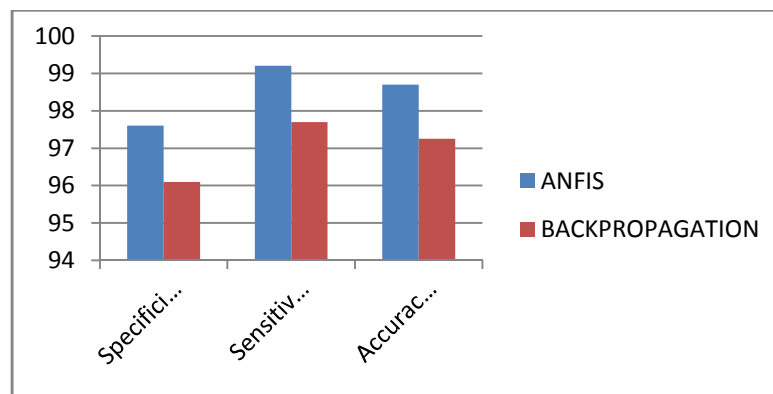


Figure 10. Comparative performance of the classifier

Impact of individual features on the detection of Glaucoma is given in Table 6. Textural features when with the shape characteristics namely rim to disc ratio and cup to disc ratio there was a good improvement in the accuracy. Graph showing the performance evaluation are shown in Figure 11.

Table 6. Performance analysis of the features

Features	Sensitivity(%)	Specificity(%)	Accuracy(%)
Cup to Disc Ratio	93.5	95.3	94
Rim to Disc area	97	96.1	96.8
First order texture	92.5	89.2	91.5
Second order texture	95.1	90.7	93.7
CDR, RDR, selective textural features	99.2	97.6	98.7

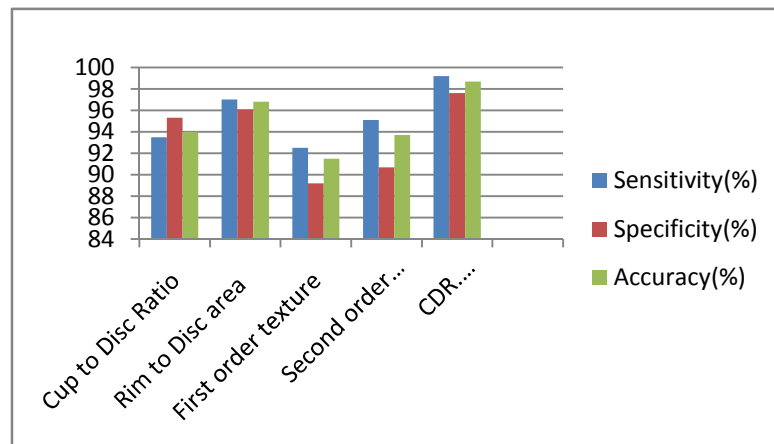


Figure 11. Performance measure of individual features.

VI. CONCLUSION

K Means clustering used in the proposed work focuses on the pallor information at each pixel thereby enabling rapid clustering and achieves a very good accuracy in detecting the optic cup. It is simple and easy to implement an unsupervised method rather than a supervised approach. Hill climbing technique and k means clustering provides a promising step for the accurate detection of optic cup boundary. Vertical CDR or superior or inferior rim area parameters may be more specific in identifying the neuroretinal rim loss along the optic disc compared to an overall cup-to-disc diameter ratio. Textural features are considered in this work in order to effectively detect glaucoma for the pathological subjects. A hybrid method involving textural features along with CDR, Neuroretinal Rim area calculation provides an efficient means to detect glaucoma. ANFIS achieves good classification accuracy with a smaller convergence time compared to Neural network classifiers. Performance of the proposed approach is comparable to human medical experts in detecting glaucoma. Proposed system combines feature extraction techniques with segmentation techniques for the diagnosis of the image as normal and abnormal. The method of considering the neuroretinal rim width for a given disc diameter with the textural features can be used as an additional feature for distinguishing between normal and glaucoma or glaucoma suspects. Progressive loss of neuroretinal rim tissue gives an accurate result to detect early stage of glaucoma with a high sensitivity and specificity. The proposed system can be integrated with the existing ophthalmologic tests, and clinical assessments in addition to other risk factors according to a determined clinical procedure and can be used in local health camps for effective screening.

ACKNOWLEDGEMENT

The authors are grateful to Dr.S.R.KrishnaDas, Chief Medical Officer and Dr. R. Kim, Chief-Vitreoretinal Service, Aravind Eye Hospital, Madurai for providing the fundus Photographs and support for our work.

REFERENCES

- [1] W.H.O, "World Health Organization Programme for the prevention of blindness and deafness-global initiative for the elimination of avoidable blindness," Document no.: WHO/PBL/97.61 Rev.1; Geneva: 1997.
- [2] Glaucoma Research foundation.(2009).[online].Available http://www.glaucoma.org/learn/glaucoma_facts.php.
- [3] C.Sinthanayothin, J.F.Boyce, C.T.Williamson, "Automated detection of the Optic disc, Fovea and retinal blood vessels from digital color fundus images," British Journal of Ophthalmology, 38, pp 902-910, 1999.
- [4] S.Kavitha, S.Karthikeyan, Dr.K.Duraiswamy, "Neuroretinal rim Quantification in Fundus Images to Detect Glaucoma", IJCSNS International Journal of Computer Science and Network Security, Vol.10, No.6, pp134-139, June 2010.

- [5] J. Liu, D. Wong, J. Lim, H. Li, N. Tan, and T. Wong, "Argali- An automatic cup-to-disc ratio measurement system for glaucoma detection and analysis framework , "Proc. SPIE, Medical Imaging, pages 72603K-8, 2009.
- [6] D. Wong, J. Liu, J. H. Lim, H. Li, X. Jia, F. Yin, and T. Wong, "Automated detection of kinks from blood vessels for optic cup segmentation in retinal images," Proc.SPIE, Medical Imaging, page 72601J, 2009.
- [7] J. Xu, O. Chutatape, E. Sung, C. Zheng, and P. Chew, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis", Pattern Recognition, 40(7):2063-2076, 2007.
- [8] Gopal Datt Joshi, Jayanthi Sivaswamy, S.R.Krishnadas, "Optic Disc and Cup Segmentation from monocular Color retinal images for Glaucoma Assessment, " IEEE Transactions on Medical Imaging,2010.
- [9] Gopal Datt Joshi, Jayanthi Sivaswamy, Kundan Karan, S. R. Krishnadas, "Optic Disk And Cup Boundary Detection Using Regional Information" Proceedings of the 2010 IEEE international conference on Biomedical imaging: from nano to Macro,2010.
- [10] J.Liu,J.H.Lim,and H.Li, "ARGALI":An automatic cup to disc ratio measurement system for Glaucoma analysis using Level set Image processing," in SPIE Medical Imaging, San Diego, USA, Feb 2008.
- [11] J. Liu, D. W .K Wong, J. H. Lim, X. Jia, F. Yin, H. Li, W. Xiong, T.Y. Wong, "Optic Cup and Disc extraction from Retinal Fundus Images for Determination of Cup- to- Disc Ratio, " in proceedings of 2008, IEEE Engineering pp 1828-1832.
- [12] Yuji Hatanaka, Atsushi Noud, Chisako Muramats, Akira Sawad, Takeshi Hara,Tetsuya Yamamoto, Hiroshi Fujita, " Vertical cup-to-disc ratio measurement for diagnosis of glaucoma on fundus images", Proc. of SPIE Vol. 7624 76243C-1,2010.
- [13] J.Nayak, U.R.Acharya, P.S.Bhat,A.Shetty and T.C.Lim, "Automated Diagnosis of glaucoma using digital fundusimages," Journal of Medical Systems,Vol.33,No.5,pp.337-346,August 2009.
- [14] K. Stapor, "Support vector clustering algorithm for identification of glaucoma in ophthalmology," bulletin of the polish academy of sciences technical sciences,"Vol. 54, No. 1, 2006.
- [15] M.Balasubramanian, S.Zabic, C.Bowd,H.W.Thompson,P.Wolenski,S.S. Iyengar, "A framework for detecting glaucomatous progression in the optic nerve head of an eye using proper orthogonal decomposition," IEEE Transactions on Information Technology in Biomedicine,Vol.13,No.5,pp.781-793,September 2009.
- [16] Mei-Ling Huang, Hsin-Yi Chen and Jian-Jun Huang, "Glaucoma detection using adaptive Neuro-fuzzy inference system, Expert Systems with Applications, Science Direct,32, 458-468, 2007.
- [17] U.Rajendra Acharya,Sumeet Dua,Xian Du,Vinitha Sree .S and Chua Kuang Chua "Automated diagnosis of glaucoma using texture and higher order spectra features, "IEEE Transactions on information technology in biomedicine,Vol.15,No.3,May 2011.
- [18] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 7, pp.629 - 639, July 1990.
- [19] T.Ohashi, Z.Aghbari, and A.Makinouch, "Hill-climbing algorithm for efficient color-based image segmentation, In IASTED International Conference On Signal Processing, Pattern Recognition, and Applications (SPPRA 2003), June 2003.
- [20] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu, "An Efficient k-Means Clustering Algorithm: Analysis and Implementation , "IEEE transactions on pattern analysis and machine intelligence, Vol. 24, No. 7, July 2002.
- [21] J.H.Tan,E.Y.K.Ng and U.R.Acharya, "Study of normal ocular thermogram using textural parameters,"Infrared Physics and Technology,Vol.53, No.2,pp.120-126,March 2010.
- [22] Jain, A., Zongker, D, "Feature selection: evaluation, application, and small sample performance," IEEE Transactions on Pattern Analysis and Machine Intelligence ,Vol.19 ,No. 2, 153–158,1997.
- [23] David. E. Goldberg, 'Genetic algorithms in search, optimization and machine learning, Addison-Wesley,pearson education ,pp59-86,1989.
- [24] J.S.R Jang, C.T Sun, E. Mizutani, Neuro-fuzzy and soft computing: approach to learning and machine intelligence, London: prentice-Hall international, vol.8, no.8. 1997.

Authors

S. Kavitha received B.E. degree in Electronics and Communication Engineering and M.E. degree in Applied Electronics from Government College of Technology, Coimbatore. She worked in Amrita Institute of Technology and Science, Coimbatore for five years. She is working as Assistant Professor in Nandha Engineering College, Erode since 2004. Her research interest includes Digital Image Processing, Neural Networks and Genetic Algorithm. She is a life member of ISTE.



K. Duraiswamy received his B.E. degree in Electrical and Electronics Engineering from P.S.G. College of Technology, Coimbatore in 1965 and M.Sc.(Engg) from P.S.G. College of Technology, Coimbatore in 1968 and Ph.D.from Anna University in 1986. From 1965 to 1966 he was in Electricity Board. From 1968 to 1970 he was working in ACCET, Karaikudi. From 1970 to 1983, he was working in Government College of Engineering Salem. From 1983 to 1995, he was with Government College of Technology, Coimbatore as Professor. From 1995 to 2005 he was working as Principal at K.S. Rangasamy College of Technology, Tiruchengode and presently he is serving as Dean of K. S. Rangasamy College of Technology, Tiruchengode, India. He is interested in Digital Image Processing, Computer Architecture and Compiler Design. He received 7 years Long Service Gold Medal for NCC. He is a life member in ISTE, Senior member in IEEE and a member of CSI.



STEP-HEIGHT MEASUREMENT OF SURFACE FUNCTIONALIZED MICROMACHINED MICROCANTILEVER USING SCANNING WHITE LIGHT INTERFEROMETRY

Anil Sudhakar Kurhekar and P. R. Apte

Deptt. of Electrical Engg., Indian Institute of Technology Bombay, Powai, Mumbai, India

ABSTRACT

Micro-cantilever arrays with different dimensions are fabricated by micromachining technique onto silicon <100> substrate. These sputtered Gold-Coated micro-cantilevers were later surface functionalized. Scanning Electron Microscopy, Atomic Force Microscopy and Optical SWLI using LASER probe are employed to characterize the morphology and image measurement of the micro-cantilever arrays, respectively. Compared with conventional AFM and SPM measurement technique, the proposed method has demonstrated sufficient flexibility and reliability. The experimental results have been analyzed and presented in this paper for MEMS Micro-cantilevers. The scanning White Light Interferometry based two point high resolution optical method is presented for characterizing Micro-cantilevers and other MEMS micro-structures. The repeatable error and the repeatable precision produced in the proposed image measurement method is nanometre confirmable. In this piece of work, we investigate the micro-structure fabrication and image measurement of Length, Width and Step-Height of micro-cantilever arrays fabricated using bulk micromachining technique onto Silicon <100> substrate.

KEYWORDS: Scanning Electron Microscopy; Atomic Force Microscopy; Micro-cantilever; Optics; Image Measurement; Silicon (100), Scanning White Light Interferometry.

I. INTRODUCTION

Step height measurement is required in many fields including semiconductors, micro-circuitry and printing. Small steps are often measured using a profilometer, calculating the least-squares straight line through the data, and then identifying the areas above and below this as being step and substrate. The step height is calculated using a least-square fit to the equation: $Z = aX + b + h \hat{I}'$ where a , b , h are unknowns and \hat{I}' takes the value of +1 in the higher regions and -1 in the lower regions. The unknowns a and b represent the slope and intercept of the line. The step height is calculated as twice the value of the third unknown, h . This approach is fine for samples where the flatness of the step and substrate both are good.

Accurate measurement of dimensions of microstructures using optical method has received much attention because of their potential advantages over conventional AFM/SPM techniques.^[1] A common method to fabricate the micro-cantilevers is to pattern the deposited continuous film using bulk or surface micromachining technique.^[1-2] However, these methods are demonstrated perfect only for sub-micron micro-cantilever arrays. As the micro-cantilever size decreases to nanometres, interesting behaviour may be expected. In particular, reduced micro-cantilever size results in change of domain structure^[2] and will affect the characterization of the micro-cantilever. One of the other methods to fabricate the nanometres micro-cantilever array is laser micromachining the deposited material onto a silicon substrate.^[2,3] Until now, the conventional image measurement technique for planar micro-structural properties of micro-cantilever on silicon <100> substrates have been studied.^[6] However, the applicability of optical methods for microstructure arrays is established. In this piece of work, we investigate the micro-structure fabrication and image measurement of Length, Width and

Step-Height of micro-cantilever arrays fabricated using bulk micromachining technique onto Silicon <100> substrate.

II. THE METHOD

Small steps are often measured using a profilometer, calculating the least-squares straight line through the data, and then identifying the areas above and below this as being step and substrate. The step height as depicted in figure 1 is calculated using a least-square fit to the equation: $Z = aX + b + h \hat{I}'$ where a , b , h are unknowns and \hat{I}' takes the value of +1 in the higher regions and -1 in the lower regions.

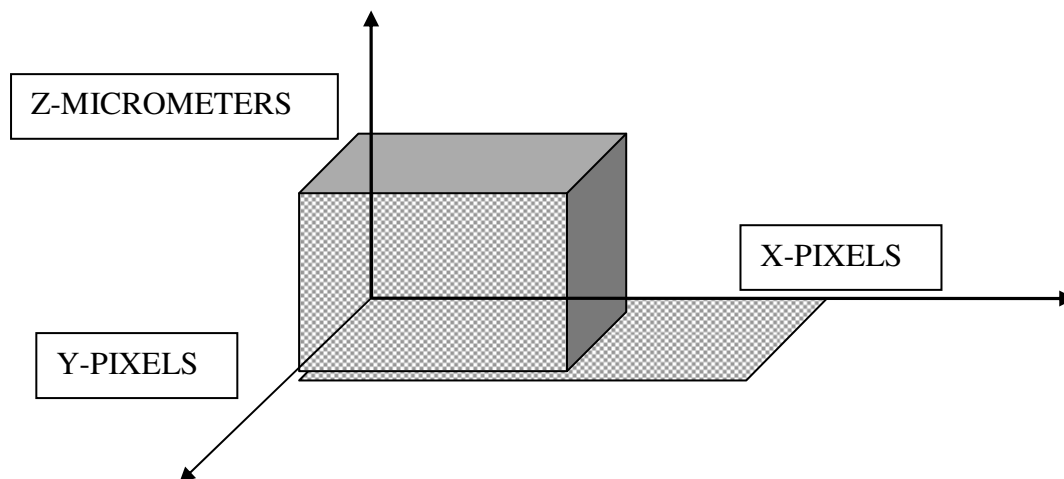


Figure 1. Three-dimensional profile of a 500-micrometer standard step height obtained from scanning white-light interferogram. The repeatability of the measurement is 10 nm.

The unknowns a and b represent the slope and intercept of the line. The step height is calculated as twice the value of the third unknown, h . This approach is fine for samples where the flatness of the step and substrate both are good.

We use optical method for precision measurements using interferometry. The ideal way to analyze complex interference data electronically is to acquire a high density of data points per interference fringe to capture all the detail of the interference signal. There are, however, practical limits to the amount of data that can be stored, processed, and displayed. This is particularly true of scanning white-light interferometry (SWLI) for surface topography measurement. These instruments use broadband sources together with mechanical translation of the object or reference surface to measure large discontinuous surface features. Typically, a SWLI instrument acquires three to five intensity values per interference fringe per pixel and process millions of data values to generate a single three-dimensional image. The volume of data involved means that 500 micrometers of depth range can require several minutes just for data acquisition. A piezoelectric actuator (PZT) is used to translate the object in a direction parallel to the optical axis of the interferometer over an interval of several tens of micrometers. The resulting interference pattern for a single pixel resembles the data simulation. The traditional way of measuring surface topography with such a system is to calculate the fringe contrast as a function of scan position and then relate the point of maximum contrast to a surface height for each pixel in the image. There are several ways to calculate the fringe contrast for this purpose, including measuring the maximum and minimum intensity values, by standard phase-shift interferometry formulas or digital filtering. These fringe-contrast techniques have in common a high density of image frames over the range of PZT translation.

The proposed method uses the fundamental physical concept of frequency domain processing of interferograms is that a complex interference pattern may be considered the incoherent superposition of several simple single-frequency interference patterns. Each of these single-frequency patterns may be represented by the simple formula,

$$I = 1 + \cos(\phi) \quad (1)$$

Where,

$$\phi = k \cdot Z \quad (2)$$

Here k is the angular wave number and the distance Z is the phase-velocity optical path difference in the interferometer. For simplicity we assume a perfectly compensated and intensity-balanced interferometer. From Eq. (2) it is clear that

$$Z = d\phi/dk \quad (3)$$

Hence one way to measure distances is to calculate the rate of change of phase with spatial frequency. To calculate this rate of change, we need phase values ϕ over a range Δk of frequencies centered around a mean frequency k_0 .

A simple linear fit to the phase data provides the rate of change $d\phi/dk$ and the mean phase ϕ_0' . This information can be used to calculate the distance in either one of two ways. The rate of change of phase can be used alone to calculate distance with Eq. (3). Alternatively, this preliminary calculation can be used to remove the 2π ambiguity in the mean phase ϕ_0' , which may then be used in an inverted form of Eq. (2) for high-precision measurements.

The Laser probe is used to select the points of interest on the device structure. The Measurement of Length and Width is relatively simple because the two-points placed shall be in-plane. However, for the measurement of Step-Height, the two-points shall not be in-plane. So, we fix one point (Maximum Z) using the marker and the other point can be placed where the height need to be measured (Minimum Z) as shown in Figure 2.

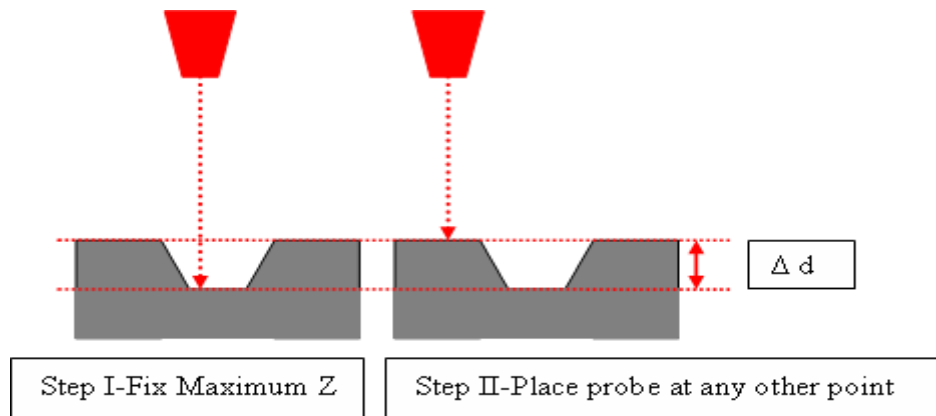


Figure 2. Proposed Two-point optical Method for Etch-Depth Measurement

The difference between Maximum Z and Minimum Z of the markers (Δd) shall give the Step-Height. The film Step-Height (Δd) is directly proportional to Wavelength of the Laser Light and is inversely proportional to twice the refractive index of the film being etched. Thus, the change in film Step-Height, using this proposed method, is measured with the relation (4),

$$Z = \Delta d = \lambda / 2 \cdot \eta \quad (4)$$

Where, λ is the wavelength of the laser light and η is refractive index of the etched layer.

III. EXPERIMENTAL

Two procedures were used to fabricate micro-cantilever arrays. Firstly, Silicon<100> substrates were deposited by silicon-di-oxide using a thermal oxidization procedure as described previously.^[4] The oxide deposition rate was about 1.5 nm/min and the gas flow rate of 18 sccm. After the oxidization and Patterning, the residual Silicon was removed by an anisotropic etchant.^[5] Secondly, micro-cantilever arrays were deposited by RF magnetron sputtering from a gold target at room temperature. The sputtering chamber was firstly pumped down to 180 milli Torr. Then, the deposition of Chrome was carried out under Ar atmosphere with about 180 milli Torr and the gas flow rate of 18 sccm. During the deposition process, the continuous film of gold was also deposited onto a silicon substrate under the same condition for the convenience of measuring the film thickness.

The sputtered chrome-gold layer has affinity with thiophenol molecules. ^[7] Considering this fact, we have dip-coated the piezo-resistive micro cantilevers with 1 micromole thio-phenol in ethanol solution for 3 Hrs. and then, rinsed with ethanol, for 2 minutes. The surface becomes functionalized.

The surface morphology of the micro-cantilever arrays was investigated by the scanning electron microscopy (SEM, JEOL 2000). The grazing incidence LASER diffraction (GILD), which avoids the effect of substrate to the pattern, was used to study the image measurement of the microstructure. The mechanical properties at the temperature 300 K were measured by Atomic Force Microscopy.

Image measurement of significant parameters of surface functionalized micro machined micro cantilevers such as Length Width and Step- Height were obtained using SEEBREZ[®] optical multi-sensing system with laser probe and Taylor-Hobson's Form Talysurf[®] 3-D surface profiler machine with 3-D Telemap Gold 3-D profile software. The Coni-spherical Stylus with the base radius of 2 micrometers was used for the contact-mode measurements.

The co-ordinate measurements were done with SEEBREZ[®] optical multi-sensing system with laser probe. This system has an auto focus facility. After the sample was prepared for measurements, the origin of the wafer co-ordinates was put relatively. Then Maximum z-coordinate was fixed with a laser-beam marker. The Measurement of Length and Width is relatively simple because the two-points placed shall be in-plane. However, for the measurement of Step-Height, the two-points shall not be in-plane. So, we fix one point (Maximum Z) using the marker and the other point can be placed where the height need to be measured (Minimum Z). The difference between Maximum Z and Minimum Z of the markers (Δd) shall give the Step-Height. The change in the film Step-Height (Δd) is directly proportional to Wavelength of the Laser Light and is inversely proportional to twice the refractive index of the film being etched. Thus, the change in film Step- Height, using this proposed method, is measured with the relation (5),

$$Z = \Delta d = \lambda / 2 \cdot \eta \quad (5)$$

Where, λ is the wavelength of the laser light and η is refractive index of the etched layer. We have measured Length, Width and Step- Height using the proposed method, which is relatively easy and accurate. The x, y and z co-ordinates of any part of the structure is measured accurately.

IV. RESULTS AND DISCUSSION

Figure 2 shows the proposed two-point-high-resolution and accurate optical method for image measurement of the micro-cantilever array. Figure 3 shows the SEM picture of the micro-cantilever array with a thickness of 500 nm.

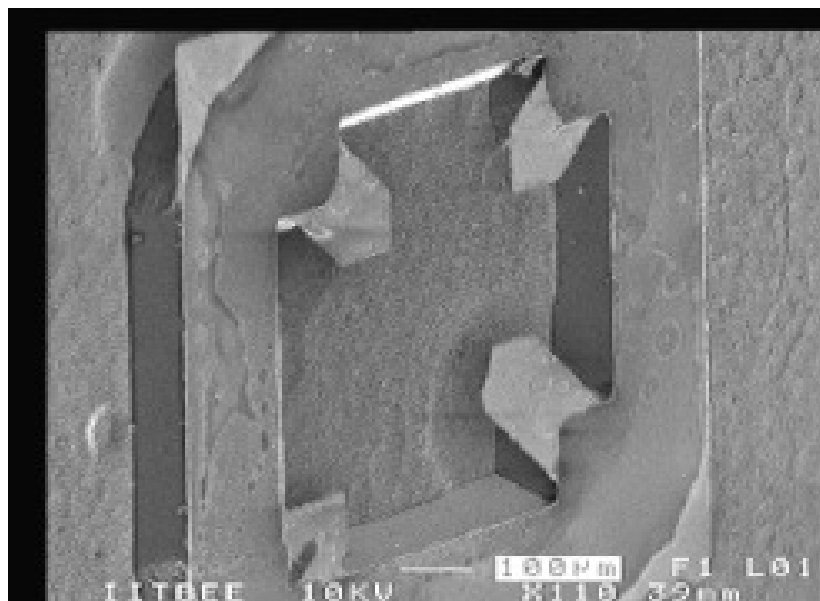


Figure 3. Scanning Electron Microscopy Micrograph of Micro cantilever on silicon<1 0 0> surface.

The micro-cantilever has a wider parameter distribution, which the mean parameters deviation is approximately 10 nm. Further analysis of SEM shows that there is some etch-product reminiscent at the bottom of the trapezoidal etch-pit. However the side-walls of the etch-pit are smooth. Figure 4 shows the Atomic Force Microscopy of the sample. It shows that the film with micro-cantilevers is free standing in a trapezoidal micro-cavity. Further analysis shows that the crystallite size distributes in the range of 5-10 nm, calculated with the Scherrer formula.



Figure 4. Atomic Force Microscopy Micrograph of Surface Functionalized Micro-cantilever on silicon<1 0 0> surface

To get the etch-depth information, the micro machined sample was kept on the base of a Taylor-Hobson's- Form Talysurf® 3-D surface profiler machine. The Coni-spherical Stylus was reset to original position after deciding the frame of x-y travel. The machine starts building etch-profile slowly as the conic-spherical stylus moves over the sample. Three-Dimensional Profile generated using Taylor Hobson's Talysurf Profiler machine of the sample are shown in Fig. 5, with the contact mode using conic-spherical stylus moving parallel to the film plane.

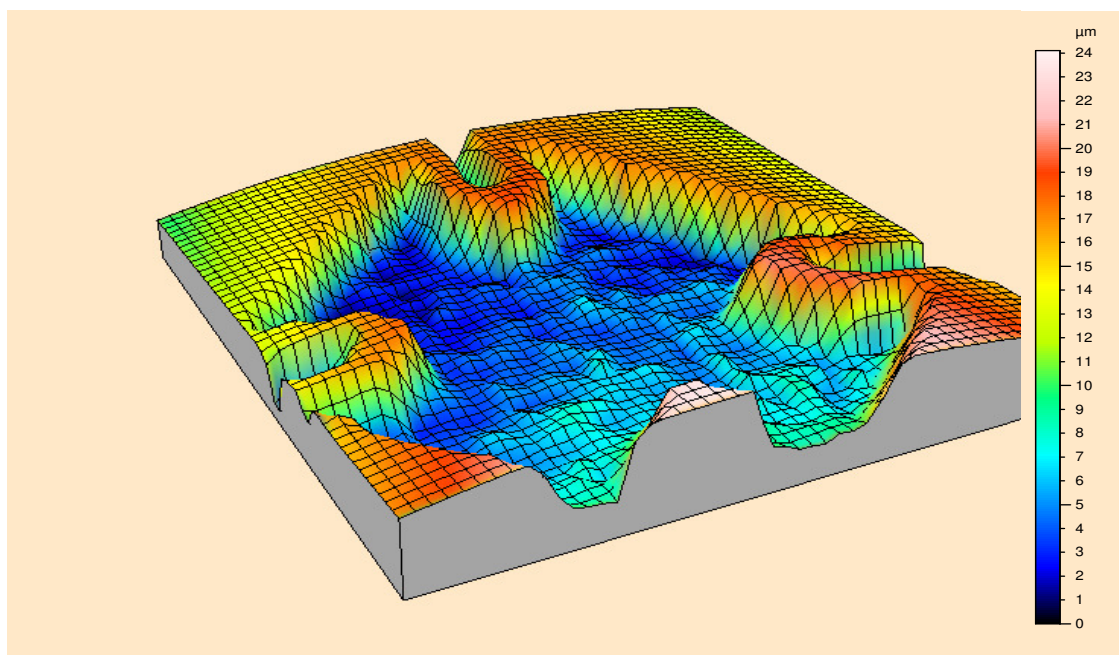


Figure 5. 3-D Surface Profiles (Silhouettes) of Micro cantilever.

The image generated using the stylus based profilometry is in conformity with the Scanning Electron Microscopy. Figure 6 and Figure 7 depicts the width and length measurement with SEEBREZ machine with high accuracy and repeatable precision.

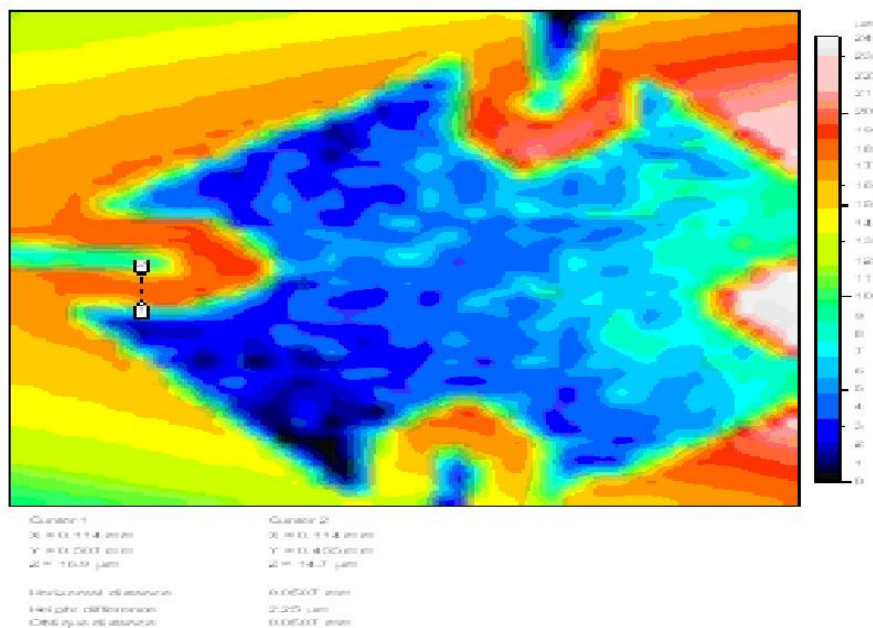


Figure 6. Width measurement of Micro cantilever.

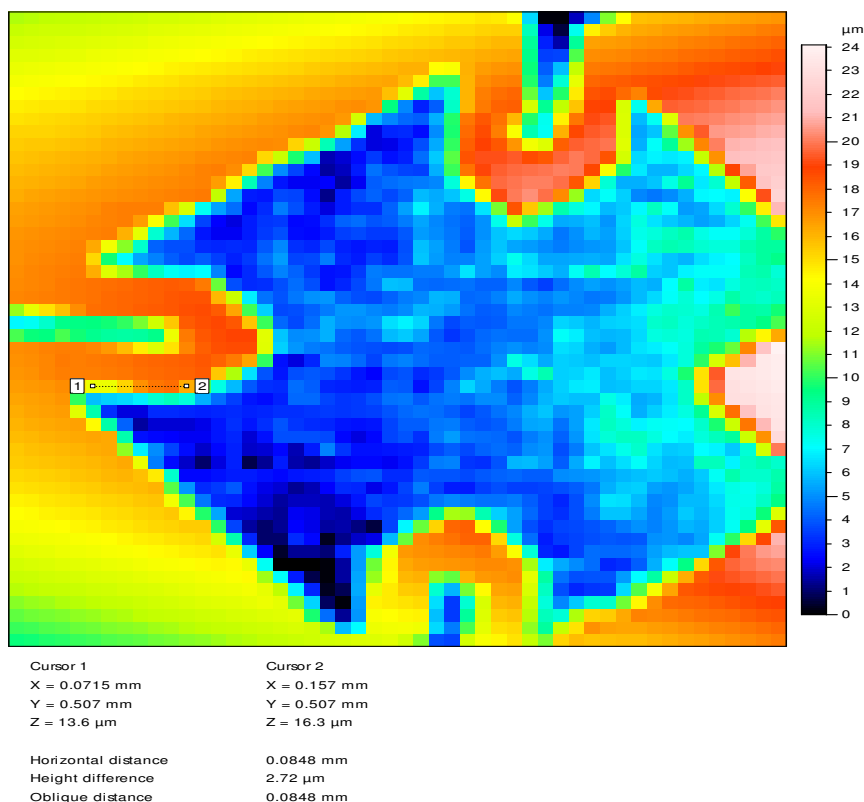


Figure 7. Length measurement of Micro cantilever.

A two-dimensional profile to recover the depth information is depicted in Figure 8.

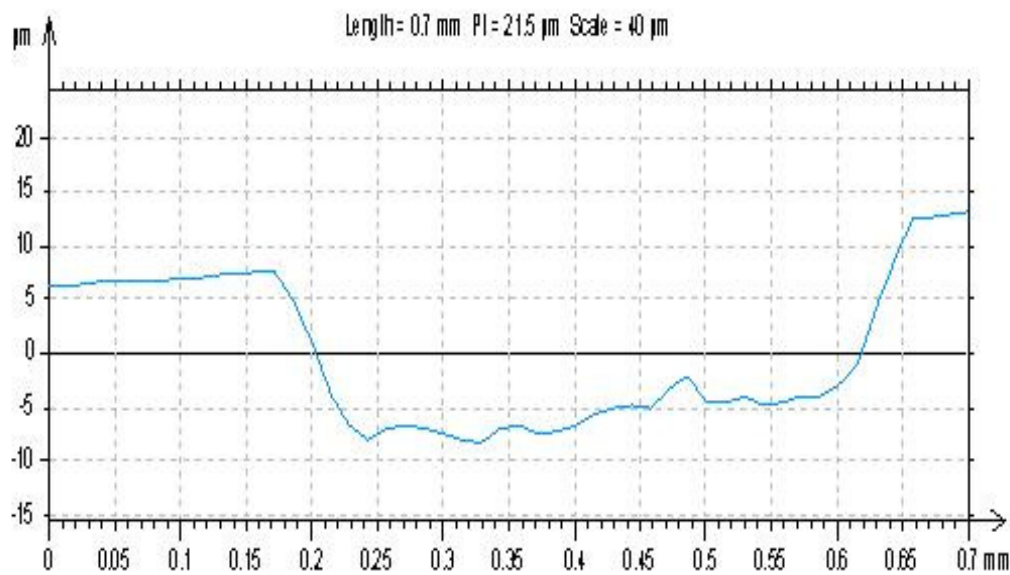


Figure 8. Etch-profile of Micro cantilever.

Table 1 depicts the designed and measured dimensions using the proposed method.

Table 1: Dimensions of Micro cantilever: Designed and Measured Using proposed Method

Micro-Cantilever			
Sr.No.	Dimensions	Designed	Measured after micro-machining
1	Length	200 micrometers	184 \pm 0.01 Micrometers
2	Width	60 micrometers	50 \pm 0.01 Micrometers
3	Step-height	200 Nanometres	180 \pm 0.01 Nanometres

V. CONCLUSION

In summary, Micro-cantilevers array have successfully been fabricated on silicon <100> substrate using bulk micromachining technique, deposited with a chrome-gold layer using the RF magnetron sputtering method. These Micro-cantilever arrays were surface functionalized using 1 micro-mole Thio phenol in ethanol solution for 3 Hours, after rinsing with ethanol solution for 2 minutes. The Micro cantilever Surface become functionalized for mass sensing. The Length, Width and Step-height measurement of micro-cantilever is obtained with the proposed high-resolution and accurate two-point optical non-contact method.

Etch profile is very important in assessing the etch-uniformity, side-wall smoothness and etch-depth. Etch-profile also infers the shape, the slope and the etch-depth of the micro cavity, in which the micro cantilevers are free-standing. It is obvious, from the inspection of etch-profile of the trapezoidal micro cavity, that the side-walls of this anisotropically etched trapezoidal micro cavity are smooth, since the profile is not jagged. However, the bottom of the anisotropically etched micro cavity is not smooth, since the profile line is jagged containing hills and dales where surface tension in the inks or paints causes either a rounding or dimpling of the "step", and the stresses caused by the curing process can cause distortion in the substrate. The curvature of the substrate might well be sufficient to prevent the use of the simple least-squares line fit. The step and substrate areas are then treated as line segments, allowing the curvature of the substrate to be removed, resulting in a straight-line representation of the substrate. The step heights are calculated from this line in the areas adjacent to each step.

The inspection of Etch-profile and Scanning Electron Micrograph confirm that the side-walls are smooth and at the bottom of the trapezoidal micro-cavity there is some etch product reminiscent. With this substantial conclusion, we propose a high-resolution accurate method for exact measurement of

Length, Width and Etch-Depth of the micro-machined micro-cantilever. Further, this method can be extended for the measurement of significant parameters of other out-of-the-plane MEMS structures. The principal disadvantage with use of under sampled data is of course a reduced signal-to-noise ratio not only because fewer data are acquired but also because of aliasing of noise at higher frequencies than the sample rate. However, this reduction in signal-to-noise ratio may in many circumstances be offset by the ability to average several scans during a shorter period and with less computer storage than is required by conventional methods. The rapid scans also reduce the sensitivity of the instrument to certain environmental effects, such as mechanical drift due to temperature and vibration.

ACKNOWLEDGEMENTS

The author acknowledges the Microelectronics Group, Nanoelectronics Centre, Department of Electrical Engineering, Department of Chemistry, Department of Physics, Suman Mashruwala Micromachining Laboratory – Department of Mechanical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai and Precision Engineering Division, Bhabha Atomic Research Centre, Trombay, Mumbai, INDIA.

REFERENCES

- [1] P. R. Apte, U.D. Vaishnav, S.G. Lokhare, V. R. Palkar, S.M. Pattalwar, "Micromechanical components with novel properties", Proc. Of SPIE, 3321, 287-297 (1996)
- [2] Marc J. Madou, "Fundamentals of Micro-fabrication: The Science of Miniaturization", CRC Press, (2002)
- [3] Xinlin Wang, "Femtosecond laser direct fabrication of metallic cantilevers for micro-corrosion-fatigue test" *J. Micromech. Microeng.* **17** 1307 (2007)
- [4] Deal Bruce, "The Oxidation of Silicon in Dry Oxygen, Wet Oxygen and Steam" *J. Electrochem.Soc.* Vol.10.issue 6, pp 527-533 (1963)
- [5] Elin Steinsland, Terje Finstad and Anders Hanneborg, "Etch rates of (100), (111) and (110) single-crystal silicon in TMAH measured in situ by laser reflectance interferometer", *Sensors and Actuators A: Physical.* Vol.86, issue (1-2), pp.73-80 (2000)
- [6] Z Sun and A Weckenmann, "Reflective Properties of Typical Microstructures Under White light Interferometer" *Meas. Sci. Technol.* Vol.22, Number 08, 103 (2011)
- [7] John A. Seelenbinder, Brown Chris W and Urish Daniel W, "Self-Assembled Monolayers of Thiophenol on Gold as a Novel Substrate for Surface-Enhanced Infrared Absorption", *Applied Spectroscopy.* Vol. 54, Issue 3, pp. 366-370 (2000).
- [8] F. Rémacle, E.S. Kryachko, "Thiophenol and thiophenol radical and their complexes with gold clusters Au₅ and Au₆", *Journal of Molecular Structure*, Volume 708, Issues 1-3, 1 December 2004, Pages 165-173, ISSN 0022-2860, 10.1016/j.molstruc.2004.02.056.

Authors

A. S. Kurhekar was born in India, in Year 1966. He received the Bachelor degree from the Amaravati University of India, in Year 1988 and the Master in Engineering degree from the Dr. B. A. M. University of India, in Year 1993, both in Electronics engineering. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology, Bombay, India. His research interests include MEMS Layout, Design, Simulation and Fabrication.



P. R. Apte was born in India, in Year 1947. He received the Bachelor degree from the Indore University of India, in Year 1968 and the Master of Technology degree from Indian Institute of Technology, Kanpur, India, in Year 1993, both in Electronics engineering. He was conferred Ph.D. by University of Mumbai in 1988. He is currently a Professor with the Department of Electrical Engineering, Indian Institute of Technology, Bombay, India. He was a member of the team that made the first TTL IC in India in 1972. Since then, He has more than 20 years experience in MOS/Bipolar IC design and fabrication. He has worked at Stanford IC labs for 2 years (1977-78) as a visiting research associate. To his credit, there are 79 publications in the Silicon Technology, which includes Journal and International conference papers. His current research interest includes MEMS - Design, Layout, Mask-making, Fabrication, Packaging, Testing, Reliability etc.



EXPERIMENTAL INVESTIGATION ON FOUR STROKE CERAMIC HEATER SURFACE IGNITION C.I. ENGINE USING DIFFERENT BLENDS OF ETHYL ALCOHOL

R.Rama Udaya Marthandan¹, N.Sivakumar², B. Durga Prasad³

¹Research scholar, Dept. of Mech. Engg., Sathyabama University, India

²Asst. Professor, Sun College of Engineering and Technology, Kanyakumari Dist., India

³Associate professor, JNTU, Ananthapur, A.P., India

ABSTRACT

In this paper an experimental investigation on the performance of surface ignition ceramic heater four stroke CI engine fueled with pure diesel (B0D100E0) and ethanol-diesel blends containing 10%, 20%, 25% and 30% by volume of ethanol are evaluated. n-butanol (B) additive is used to solubility of ethanol (E) in diesel (D), that acts as a bridging agent through molecular compatibility and bonding to produce a homogeneous blend. The ethanol – diesel fuel affects blend stability, viscosity, lubricity, corrosiveness and safety. The tests are carried out on 10HP ceramic heater surface ignition single cylinder diesel engine under steady state operating conditions. The engine is run at various speeds of 1250rpm and 1500 rpm. The relevant parameters such as brake thermal efficiency (BTE), brake specific fuel consumption (BSFC) and emissions are calculated for pure diesel and ethanol-diesel blends by B5D85E10, B5D75E20, B5D70E25 and B5D65E30. The Partially Stabilized Zirconia (PSZ) ceramic heater is used to reduce the emissions by 220 ppm of NO_x, under half load for the blends of B5D85E10 gives minimum CO emissions and unburned HC emissions by 24 ppm from the engine and improve engine output behavior to 2%.

KEYWORDS: ethanol, n-butanol, emissions, ceramic heater.

I. INTRODUCTION

Ethanol is one of the possible fuel for diesel replacement in CI engines [1]. It can be made from raw materials such as sugarcane, sorghum, corn, barley, cassava, sugar beets etc. A biomass-based renewable fuel, ethanol has cleaner burning characteristics, and a high octane rating. The application of ethanol as a supplementary compression-ignition fuel may reduce environmental pollution, strengthen agricultural economy, create job opportunities, reduce diesel fuel requirements and thus contribute in conserving a major commercial energy source [2].

A surface ignition ceramic heater [3] CI engine is able to operate at higher temperature enabling combustion of fuel at complete resulting to increase combustion efficiency. This should increase engine performance, decrease fuel consumption and reduce pollution [4]. Ceramic heater provides instant heat within seconds of turning, which helps save fuel and reduce emissions. It is mounted through the engine head, that heats up and warms air moved over its surface, due to its inherent self-regulating characteristics. Ceramic heater for diesel combustion would represent a simple low cost and easy approach in diesel engine performance [5].

In the C.I engines a premixed fuel air vapor is drawn in during the suction stroke, a single high intense spark passes across the electrode, producing a core of flame from which the combustion spreads to the envelope of mixture surrounding it at a fast rate. The above two methods evidently show that the fuel properties of the first method will not be suitable for the second, and hence if we need to have an engine with multi fuel capability, the nature of combustion should be very different from the above

methods. This is where the concept of surface ignition comes into active consideration. Surface ignition indicates the beginning of combustion from a hot surface. It will be interesting to know that almost all fuels exhibit this property to varying degrees, the alcohols being highly susceptible to this kind of combustion [8].

II. EXPERIMENTAL WORK

A different speeds stationary four stroke surface ignition[8] ceramic heater[9] CI engine is selected for the experiment. The major specifications of the engine are given in Table 1 and properties of fuel used are given in Table 2. The engine is connected with dynamometer, air stabilizing tank, diesel and ethanol blends consumption measuring device, exhaust gas analyzer etc[6]. A ceramic heater is fixed inside the cylinder and connected by 12 volt DC battery to heating combustion chamber. Diesel fuel and ethanol-diesel blends with additive by B0D100E0, B5D85E10, B5D75E20, B5D70E25 and B5D65E30 are tested. The ethanol and additive is obtained from the local market. The engine is run on no-load condition and its speeds are adjusted to 1250 rpm and 1500 rpm by adjusting the screw provided with the fuel injector pump. The engine is run to gain uniform speed after which it is gradually loaded. The experiments are conducted at six power levels for each load condition. The engine is run for at least 7 minutes after which data is collected. The experiment is repeated 5 times and the average value is taken. The observations are made during the test for the determination of various engine parameters like Brake specific fuel consumption, Brake thermal efficiency and exhaust emissions[7].

Heat transfer affects engine performance, efficiency, and emissions. The mass of fuel within the cylinder, higher heat transfer to the combustion chamber walls, will lower the average combustion gas temperature and pressure, and reduce the work per cycle transferred to the piston. Thus specific power and efficiency are affected by the magnitude of engine heat transfer [8]. Advances in engine technology by introducing ceramic heater increase the engine output efficiency and reduce the emission parameters[10].

In ceramic heater C.I engine an injection pressure and rate of injection can also offset the adverse effect of ceramic heater as shown in Figure 1. In this new system, decrease in pre mix of combustion due to decrease in ignition delay increases the Brake Specific Fuel Consumption (BSFC). Partially Stabilized Zirconia(PSZ) Ceramic Heater is fitted inside the cylinder, because of its very high fracture toughness among ceramics, it has one of the highest maximum service temperatures (2000°C) among all of the ceramics and it retains some of its mechanical strength close to its melting point (2750°C). PSZ ceramic heater is used in diesel engine because of two very notable properties: one is high temperature capability and other is low thermal conductivity. None of the other ceramics possess a thermal conductivity as low as the zirconia. This means that engine using zirconia ceramic heater would retain much of the heat generated in the combustion chamber instead of losing it to the surroundings [9].

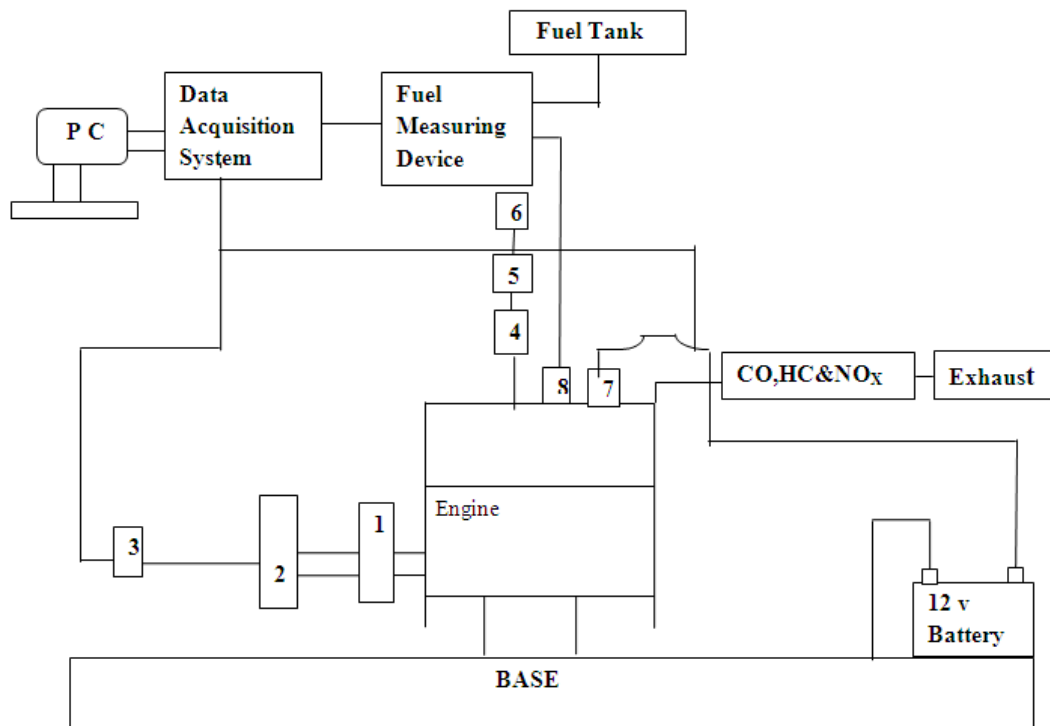
Table 1 Tested engine specifications

Engine type	4-stroke single cylinder engine
Make	Kirloskar
Power	10.0KW
Bore x Stroke(mm)	102x110
Cubic Capacity(cc)	898
Compression ration	18:1
Colling system	Water cooled
Lubrication system	Force feed
Attachment	Ceramic heater(12 V,DC)

Table 2 Properties of blending stocks

Properties	Diesel	Ethanol	n-Butanol
Boiling point ($^{\circ}\text{C}$)	180	78	117
Flash point ($^{\circ}\text{C}$)	65	10	35
Density, g/ml at 20°C	0.829	0.789	0.81
Oxygenate (Wt%)	0.84	35	7.5
Carbonate (Wt%)	-	52	25
Hydrogen (wt%)	87	13	74
Viscosity, CS at 40°C	13	1.2	10.3
Cetane number	48	6	40

Figure 1 shows the experimental set up of the work. A dynamometer is used for measuring the power of the engine output. Exhaust gas analyzer is used for measuring the emissions of CO, HC and NO_x from the engine. A fuel consumption meter is used for measuring the break specific fuel consumption of the engine, also Data Acquisition System(DAS) is used to calculate all required out put parameters.

**Figure 1** Experimental Setup

- | | |
|----------------------------|---------------------------|
| 1. Flywheel | 5. Digital air flow meter |
| 2. Dynamometer | 6. Air filter |
| 3. R.P.M. Measuring device | 7. Ceramic heater |
| 4. Air stabilizing tank | 8. Injector |

III. RESULTS AND DISCUSSIONS

The experimental tests were carried out on the surface ignition ceramic heater four stroke CI engine using pure diesel and ethanol- diesel blends with n-butanol additive at different speeds. The relevant parameters such as engine torque and fuel consumption of the engine were recorded and the brake specific fuel consumption, brake thermal efficiency were also calculated at 1250 rpm and 1500 rpm.

The engine emissions of CO, unburned HC and NO_x were analyzed using the exhaust gas analyzer. The results were obtained by data acquisition system and are shown as follows. The ratio of ethanol-diesel blends gives various fuel consumption according to the percentage of ethanol present in the diesel fuel. If more ethanol is added with diesel, gives more fuel consumption.

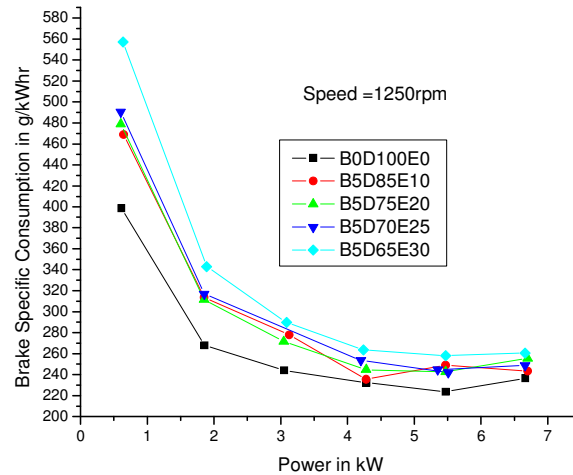


Figure 2 BSFC of the engine for 1250rpm

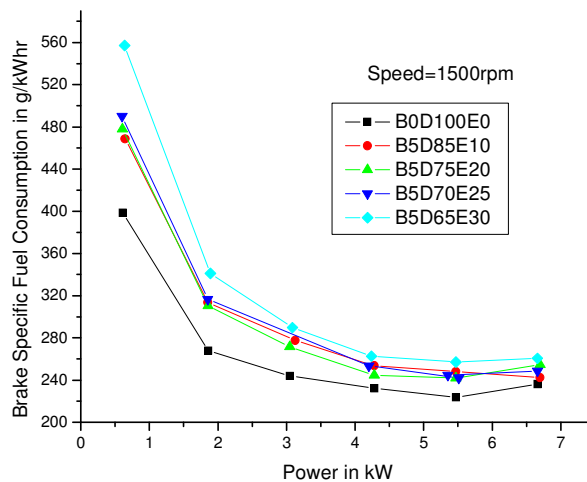


Figure 3 BSFC of the engine for 1500 rpm

Figures 2 and 3 show the BSFC of the engine. When the engine runs at 1250 rpm on different engine loads, for the blends of B5D65E30, the BSFC is increased by 4% for the blends of B5D85E10 BSFC is decreased by 1.2% for maximum engine load and the blends of B5D75E20 BSFC is average by 2.5% up and down. The results show the trends of the increase of fuel consumption with the increase percentage of ethanol in the blends.

When the engine runs at 1500 rpm, for the blends of B5D65E30, the BSFC is increased at all engine load conditions. The blends B5D70E25 give 3.25% less BSFC next to pure diesel fuel. However, BSFC is high at minimum power for all fuel ratios. This increase of fuel consumption is due to the lower heating value of ethanol than that of pure diesel [4].

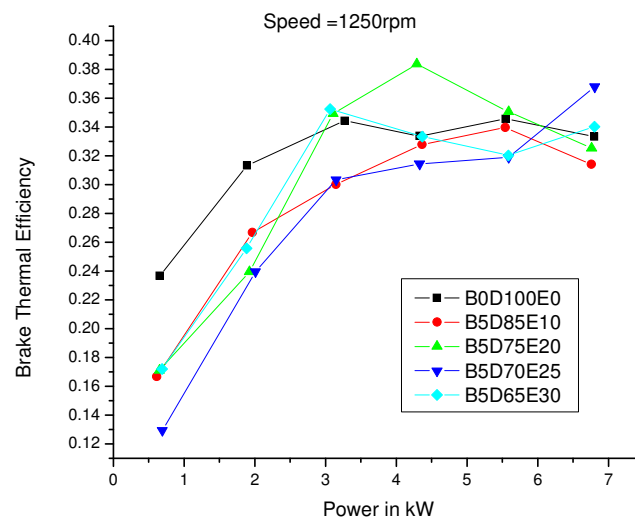


Figure 4 BTE of the engine for 1250 rpm

Figures 4 and 5 show the results of the break thermal efficiency(BTE) of the engine. When the engine runs at the speed of 1250 rpm for the blends of B5D70E25 the BTE is increased by 2.5%, and at average load BTE is increased for the blends of B5D65E30 by 3.4%. These results show the difference of the break thermal efficiencies between the blends and diesel are relatively small at 1500rpm. When the engine runs at the speed of 1500 rpm the break thermal efficiency is increased for the blends of B5D75E20 and B5D70E25 by 5% at high load and 2.5% at low load. The exhaust emissions are measured in terms of Carbon Monoxide (CO), Hydrocarbons(HC) and Oxides of Nitrogen (NO_x) emissions. The results for diesel fuel as well as ethanol-diesel blends are given below. The oxygen content of the blended fuels would help to increase the oxygen to fuel ratio in the fuel at rich regions. The resulting more complete combustion leads to reduction of CO in the exhaust. If the percentage of ethanol in the blends increased, NO_x emission is reduced. This is because of the air-fuel ratio in the case of ethanol-diesel blends, is lower as compared to diesel alone. The latent heat of vaporization of ethanol lowers at same temperature resulting in lower NO_x emissions[7].

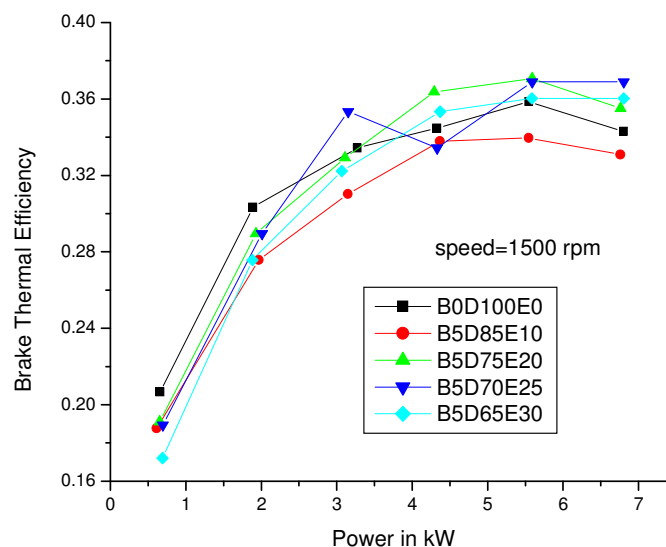


Figure 5 BTE of the engine for 1500 rpm

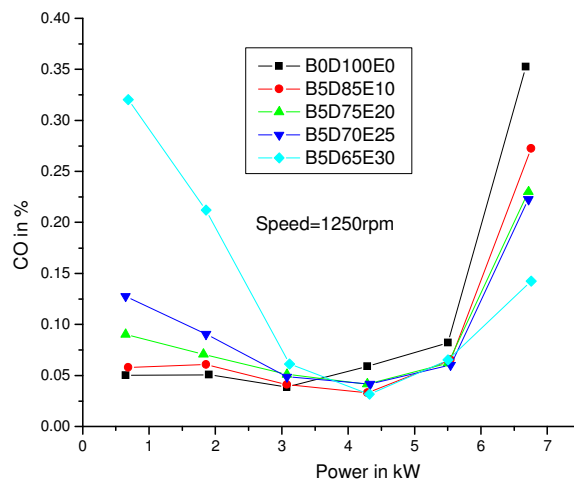


Figure 6 CO emissions for 1250 rpm

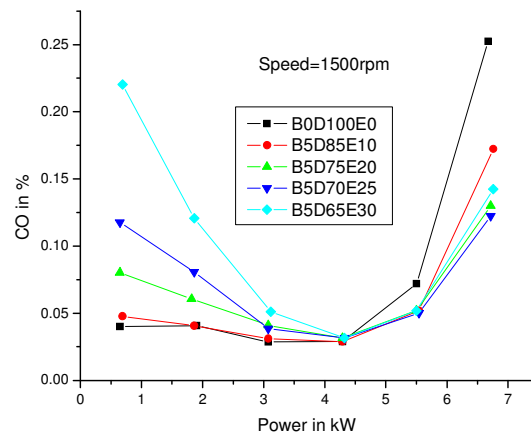


Figure 7 CO emissions for 1500 rpm

Figures 6 and 7 show the CO emissions of the engine. The CO emissions from the engine fuelled by the blends are higher than those fuelled by pure diesel at minimum loads. The higher percentage of the ethanol gives more CO emissions upto half load. For higher engine loads which are above half load, the CO emissions became lower than that fuelled by diesel for all the blends for all speeds by 0.01% to 0.07%. Upto half load the variation of CO emissions between pure diesel and B5D85E10 is only 0.05% for the speed of 1500rpm. At average load CO emission is same for 1500rpm and 0.01% to 0.02% different for 1250rpm. The percentage of ethanol in the blends increased the percentage of CO emission reduced. The emission reduced with the use of 10%, 20%, 25% and 30% ethanol-diesel blends as compared to diesel alone. This is due to the concept that ethanol has less carbon than diesel. The same fuel dispersion pattern as for diesel, the oxygen content of the blended fuels would help to increase the oxygen in fuel ratio in the fuel rich regions. This results in more complete combustion which leads to reduced CO in the exhaust smoke by ceramic heater engine.

The reduction of CO emissions at full load is due to the more complete combustion. The phenomenon is due to the oxygen element contained in ethanol. When the engine working above its half load, the temperature in the cylinder is high, which makes the chemical reaction of fuel with oxygen easier and the combustion becomes more complete for two different speeds.

The level of unburned hydro carbons (HC) in exhaust gases is generally specified in terms of the total hydro carbon concentration expressed in parts per million(ppm) carbon atoms.

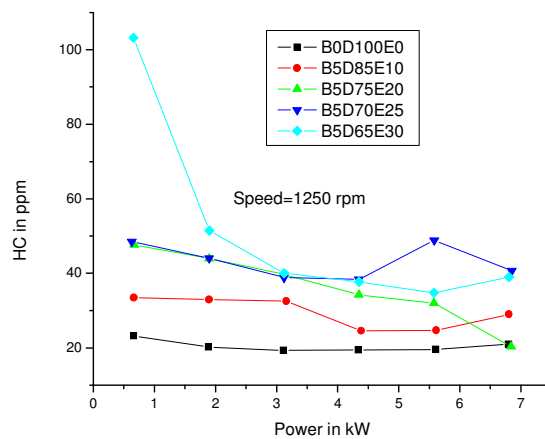


Figure 8 HC emissions for 1250rpm

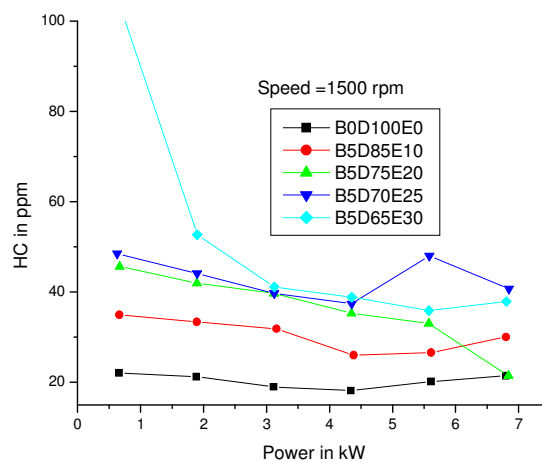


Figure 9 HC emissions for 1500rpm

Figures 8 and 9 show the HC emissions of the engine. More percentage of ethanol gives more HC emissions for all speeds. For the blends of B5D85E10 the HC is reduced by 5 to 10 ppm from pure diesel at 1250 rpm and 7 to 12.5 ppm at 1500 rpm. This is due to the high temperature in the ceramic heater engine cylinder to make the fuel be easier to react with oxygen when the engine runs on the top load and high speed. Figure 9 shows that the results of unburned HC emissions from the engine for the blended fuels are all higher when the engine runs at the speed of 1500 rpm.

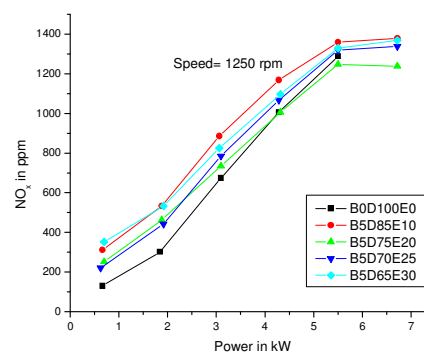


Figure 10 NOx emissions for 1250rpm

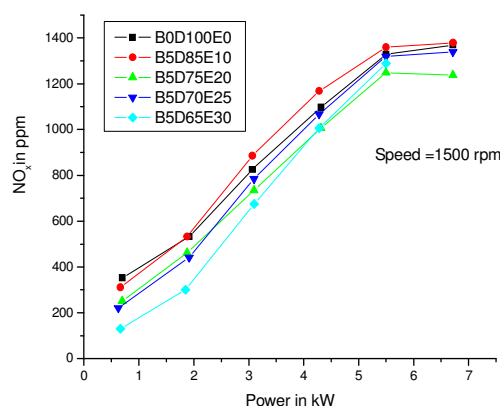


Figure 11 NOx emissions for 1500rpm

Figures 10 and 11 show the NOx emissions of the engine. When the engine runs at the speed of 1250 rpm NOx emission is minimum for pure diesel and more for 1500rpm. The NOx emission is more for the blends of B5D85E10 at 1500rpm by 20%. The blends B5D70E25 give average NOx for all working conditions. The NOx emissions from the engine are higher than those of diesel. NOx emissions are increased for all blends and speeds when the load is increased. For low load NOx is minimum due to the fuel air mixture with spread in composition about stoichiometric burns. During the mixing controlled combustion phase the burning mixture is likely to be closer to stoichiometric by the help of ceramic heater. When the engine runs at 1500rpm the NOx is reduced by 40% for the blends of B5D75E20. This is because the air-fuel ratio in the case of ethanol-diesel blends is lower as compared to diesel alone. The latent heat of vaporization of ethanol is minimum for the same temperature in minimum NOx emissions[7].

IV. CONCLUSIONS

An experimental investigation was conducted on the blends of ethanol - diesel fuel using ceramic heater surface ignition single cylinder CI engine. The tested blends were from 10% to 30% of ethanol by volume and also with 5% of the additive of n-butanol. The engine was operated with each blend at different power and different speeds of 1250 rpm and 1500 rpm.

The Experiment showed that the n-butanol was a good additive for mixing diesel with ethanol blends.

Using ceramic heater improved engine performance to 2% and controlled the emissions and reduced unburned HC by 7.5ppm.

The brake specific fuel consumption is slightly increased by 62g/Kwhr for B5D65E30 and 58g/Kwhr for B5D70E25 blends at 1250 rpm and 60g/Kwhr for B5D65E30 at 1500 rpm.

The break thermal efficiency is increased for the blends of B5D75E20 by 2% and B5D65E30 by 2.5% at the speed of 1250 rpm. When the engine runs at 1500 rpm the break thermal efficiency is increased for the blends of B5D75E20 by 5% at high power and B5D70E25 by 2.5% at low power.

The higher percentages of the ethanol give more CO emission by 70.34% maximum. At half load CO emission is average. For higher engine loads which are above half of the engine load, the CO emission becomes lower than that fuelled by diesel for all the blends for the speed of 1250rpm. CO emission is same for all blends at half load and 5% CO emission is increased at low and high load at the speed of 1500 rpm.

The blends of B5D85E10 the HC emission is reduced by 20ppm. When the engine runs at the speed of 1500 rpm, the HC emission becomes less as the loads increased. Less emission for B0D100E0 and B5D85E10 by 24 ppm and 10 ppm respectively.

NOx emission is reduced for the blends of B0D100E0 and B5D70E25 by 100ppm at the speed of 1250 rpm. When the engine runs at 1500 rpm the NOx is reduced for the blends of B5D65E30 upto half load of the engine by 220 ppm.

REFERENCES

- [1]. Alan C.Hansen, Qin Zhang, Peter W.L.Lyne, Ethanol –diesel fuel blends – a review . Bioresource Technology 96 (2005) 277 -285.
- [2]. E.A.Ajav, Bachchan Singh, T.K.Bhattacharya, Experimental study of some performance parameters of a constant speed stationary diesel engine using ethanol-diesel blends as fuel. Biomass and Bioenergy 17 (1999) 357-365.
- [3]. DANIEL NG., National Aeronautics and space Administration, Lewis Research center, Cleveland, Ohio 44135, Temperature measurement of a Miniature Ceramic heater in the presence of an Extended Interfering background radiation sources using a multiwavelength pyrometer.
- [4]. Chonglin Song, Zhuang Zhao, Gang L.V, Jinou Song, Lidong Liu, Ruifen Zhao, Carbonyl compound emissions from a heavy –duty diesel engine fueled with diesel fuel and ethanol –diesel blend. Chemosphere 79 (2010) 1033-1039.
- [5]. Douglas J.Ball and Glenn E.Tripp (Delphi Automotive), Louis S. Socha, Achim Heibel and Medha Kulkarni, Phillip A. Weber, Douglas G. Linden, A comparison of Emissions and Flow Restriction of Thin wall Ceramic Substrates for Low Emission Vehicles. 199-01-0271 Society of Automotive Engineers, Inc.
- [6]. Hwanam Kim, Byungchul Choi, Effect of ethanol-diesel blend fuels on emission and particle size distribution in a common – rail direct injection diesel engine with warm-up catalytic converter. Renewable Energy 33 (2008) 2222-2228.
- [7]. Jincheng Huang, Yaodong Wang, Shuangding Li, Anthony P.Roskily, Hongdong Yu, Huifen Li, Experimental Investigation on the performance and emissions of a diesel engine fuelled with ethanol – diesel blends. Applied Thermal Engineering 29(2009) 2484-2490.
- [8]. John B. Heywood, Internal Combustion Engine Fundamentals, Mc Graw Hill Book Company, New Delhi, 1988.
- [9]. <http://global.kyocera.com/prdct/fc/product/pdf/heaters.pdf>
- [10]. P. Satge Zde Caro, Z. Mouloungui, G. Vaitilingom, J. Ch. Berge, “Interest of combining an additive with diesel-ethanol blends for use in diesel engines”, Fuel 80(2001) 565- 574, Elsevier.

Authors

R. Rama Udaya Marthandan is working as a Professor and HOD at Sun College of Engineering and Technology, Kanyakumari District, INDIA. He received the M.E. degree in Mechanical Engineering from Annamalai University in 2000. He is a Head of the Department of Mech. Engg. He is a life member of ASME and ISTE. His research interests include Alternative Fuels and IC Engine.



N. Sivakumar is working as an Assistant Professor at Sun College of Engineering and Technology, Kanyakumari District, INDIA. He received the M.E. degree in Mechanical Engineering from Anna University: Chennai in 2004. He is a life member of ISTE. His research interests include Finite Element Analysis.



B. Durgaprasad is working as an Associate Professor at JN Technical University, Ananthapur, Andhra Pradesh, INDIA. He received his Ph.D. in the area of surface ignition IC engine from JNTU in 2000.

PERFORMANCE VERIFICATION OF DC-DC BUCK CONVERTER USING SLIDING MODE CONTROLLER FOR COMPARISON WITH THE EXISTING CONTROLLERS - A THEORETICAL APPROACH

Shelgaonkar (Bindu) Arti Kamalakar, N. R. Kulkarni
Modren College of Engg. Pune ,Maharashtra.

ABSTRACT

In recent electronic applications the variable DC power supply is derived with light weight, occupying less size using 100 kHz switching frequency. When the frequency is high, the load experiences practically uninterrupted DC voltage. According to need of application buck converter is considered for analysis. It is observed that nature of DC- DC converter is nonlinear and time variant systems, and does not lend them to the application of linear control theory. The performance of buck converter has been studied and is undertaken for their theoretical verification, graphical representation and Matlab simulation. From the linear controller PI, PID is considered and non linear controller sliding mode control is taken as control method. The paper work will highlights nonlinear aspects of buck converter, non linear controller like sliding mode controller and hybrid type of controller SMC PID. This will also focuses the benefits of non linear control.

KEYWORDS: SMC (sliding mode control), PI and PID control.

I. INTRODUCTION

DC–DC converter convert DC voltage signal from high level to low level signal or it can be vise versa depending on the type of converter used in system. Buck converter is one of the most important components of circuit it converts voltage signal from high DC signal to low voltage. In buck converter, a high speed switching devices are placed and the better efficiency of power conversion with the steady state can be achieved. In this paper work performance of buck converter is analyzed. The circuit may consist of nonlinearity like delay, hysteresis etc. and because of this output voltage is not constant. To settle the output voltage within minimum settling time and less overshoot different types of controllers are considered such as linear controller PI, PID and in nonlinear controllers SMC (sliding mode controller).The paper deals with comparison of performance of DC-DC buck converter using controllers PI, PID, SMC and SMC PID. The performance of buck converter has been analyzed in many papers amongst them papers [1][2] have been studied and are undertaken for their theoretical verification, graphical representation and Matlab simulation.

II. SIMULATED MODEL OF BUCK CONVERTER

Simulated model of buck converter by using Matlab are as shown in figure no.2.1. It consist of 24 V input DC supply, GTO (gate turn off thyristor) as switch, PWM (Pulse width modulation) generator for providing switching pluses to GTO. Inductor is of 69 μH [1] and capacitor is of 220 μF [1], with load resistance of 13 Ω [1] .The desired output from this converter is 12 V DC.

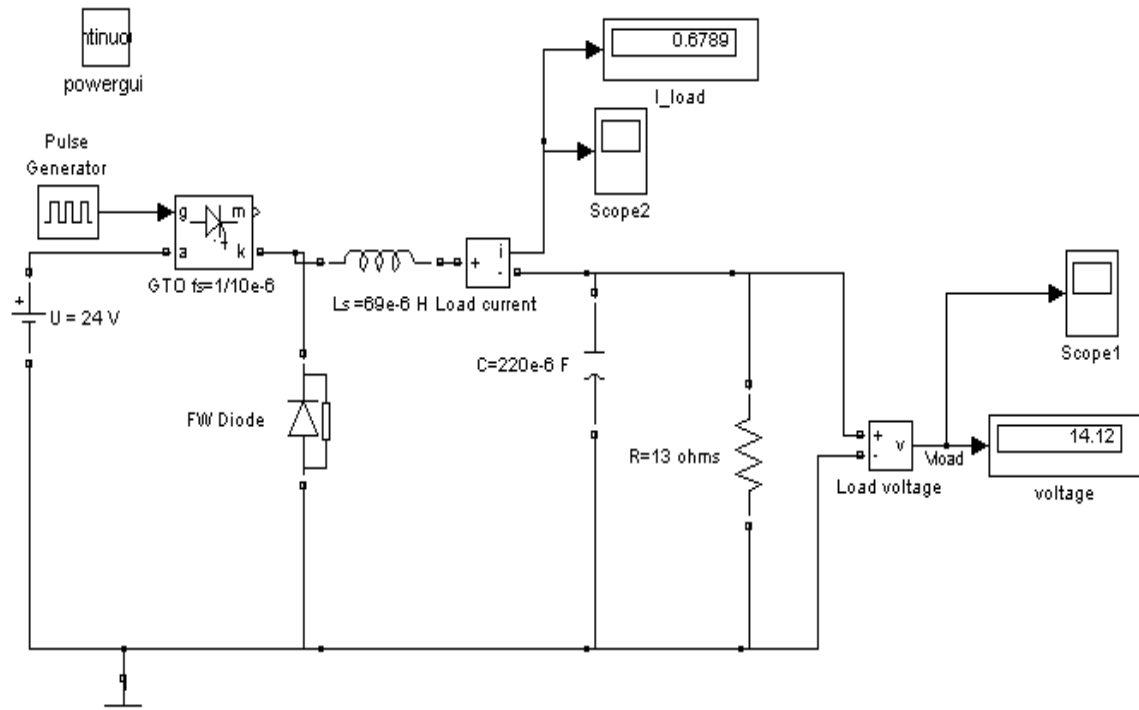


Figure No 2.1 Buck converter in Matlab Simulink.

The circuit has settling time of 2 msec and output voltage is 14.12 V which is required to settle at 12 V. To compensate these transients present in buck converter different types of controllers can be used.

III. CONTROL METHODS

Figure 3.1 shows the block diagram with some methods that can be used to control DC-DC converters and the disturbances that have influence on the behavior of the converter and its stability. The feedback signal may be the output voltage, the inductor current, or both. The feedback control can be either analog or digital control. From these control methods PI, PID are linear control methods and SMC, SMC PID are the non-linear control methods. Comparison between linear and nonlinear control methods are given below.

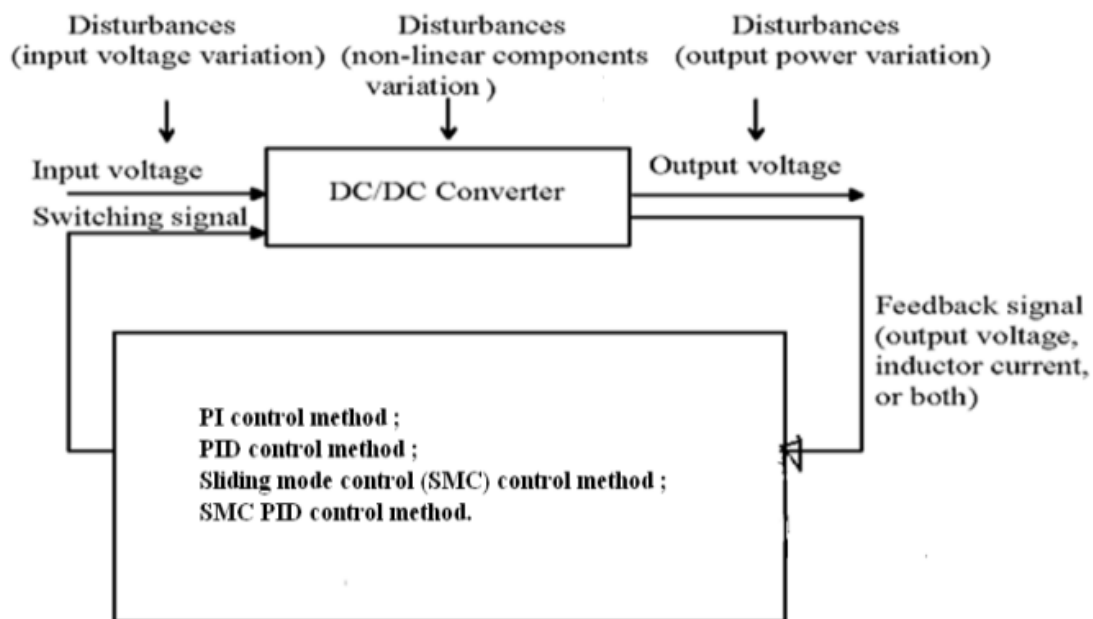


Figure No. 3.1 Types of controller.

3.1 PI control method

Settling time of PI compensated buck converter circuit is 11 msec initial overshoot for output voltage is 27 V and 43 A for inductor current. After settling time of 11 msec output voltage is at 12V and inductor current is at 1.738 A.

Load Voltage:-

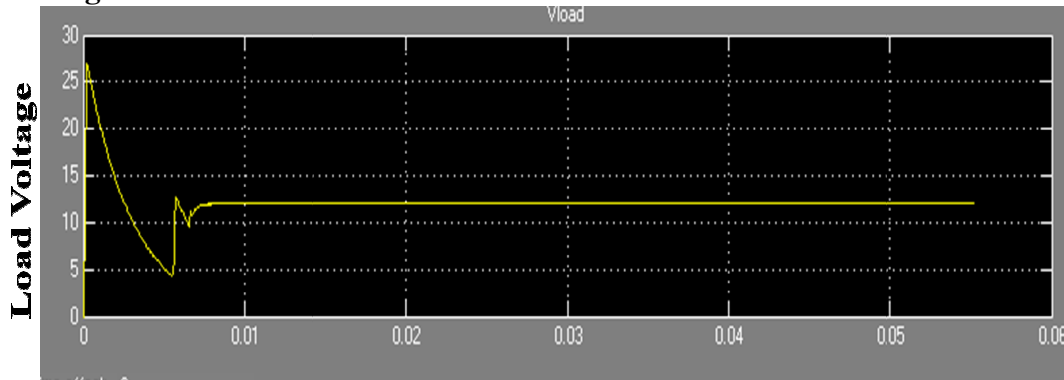
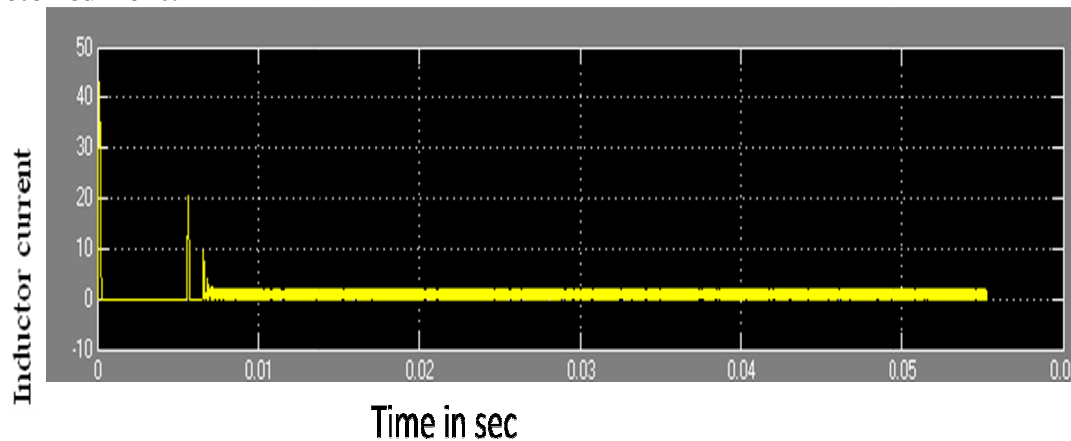


Figure No. 3.1.1 Load voltage of buck converter in Matlab/Simulink™ model.

Inductor current:-



FigureNo.3.1.2 Inductor current from simulation.

3.2. Effect of variation of load resistance on buck converter with PI control

When buck converter is considered with PI control it has settling time of 11 msec and output voltage is at 12 V. When the circuit was tested under the load variation from 0 (open circuit) to short circuit, it was found that as load resistance increases load current decreases.

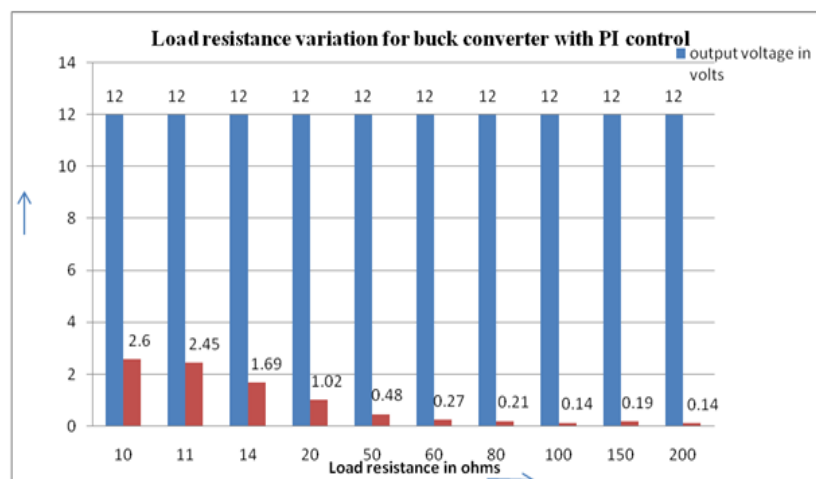


Figure No.3.2.1 Bar graph for the variation of load resistance.

3.3. Effect of variation of line voltage on buck converter with PI control

When the circuit was tested under the line variation from 20 V to 34 V, it was found that as line resistance increases, load current increases settling time is almost remains constant for PI controller.

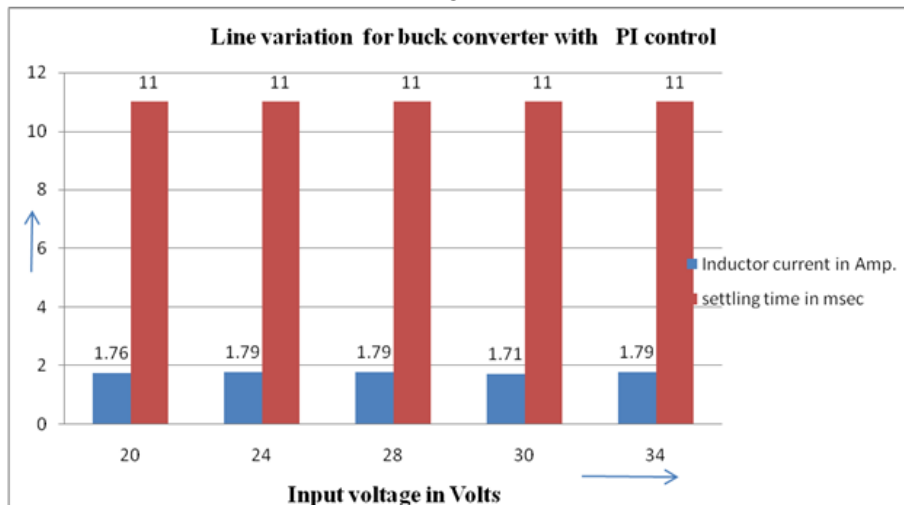


Figure No.3.3.1 Bar graph for the line variation.

3.4. PID control method

PID controllers are dominant and popular and, have been widely used since the 1940's because one can obtain the desired system responses and it can control a wide class of systems. The basic AC modeling approach is a common averaging technique used for PID modeling. After the circuit is modeled, we go through the design of PID controller with the help of Matlab in a simple way to get an overall system with good quality performance. Simulink model of the converter is built up and the controller obtained is added to the model.

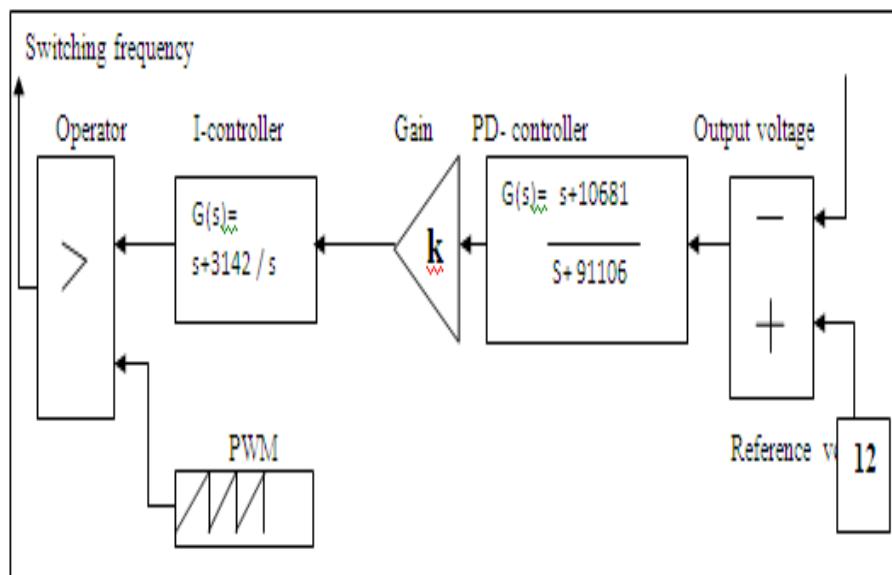


Figure No 3.4.1 The block diagram of controller includes PID control.

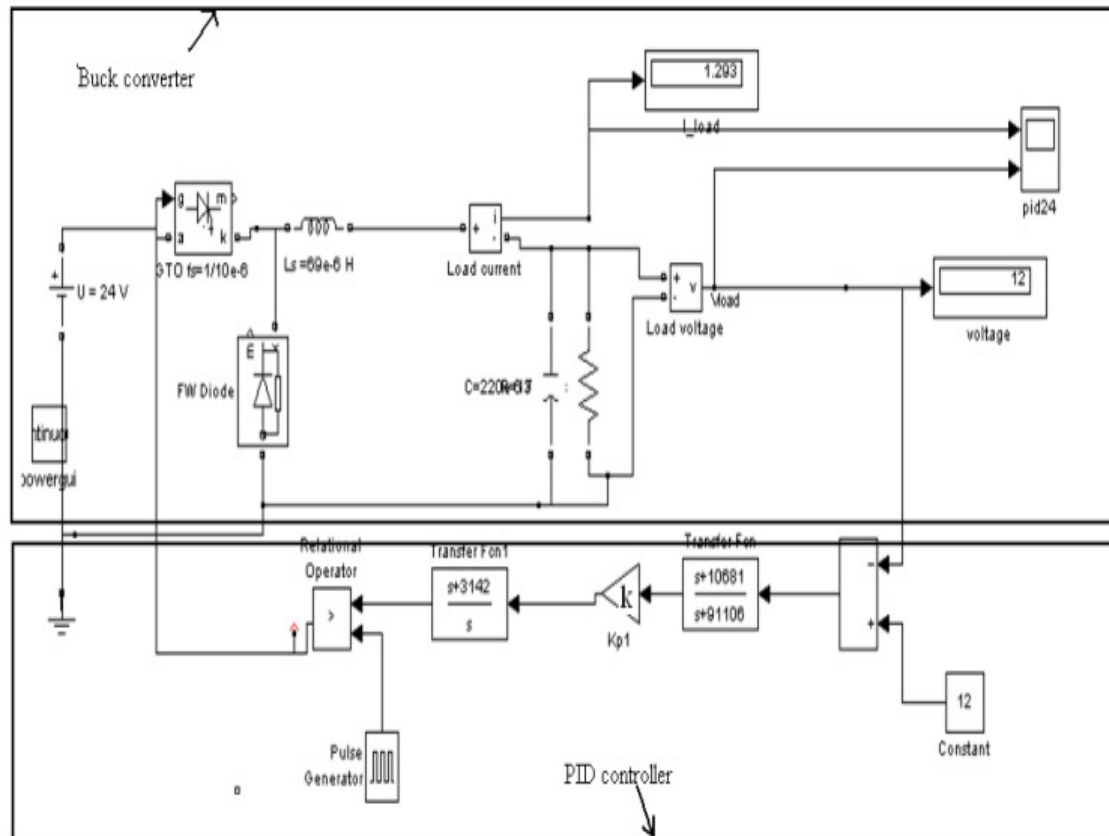


Figure No. 3.4.2 Buck converter with PID control Matlab model

3.4.1. Inductor current waveform

By considering above scenario in which a buck converter when considered with PID controller it has been observed that the circuit has settling time of 2.5 msec. The output voltage attends steady state value of 12 V, which is expected output from this application. Settling time for PID controlled buck converter is 2.5 msec and transient voltage is of 16 V and transient current is of 28 A which are less as compared to PI controller.

3.5 Effect of variation of load resistance on buck converter with PID control

When PID controlled buck converter is considered with load variation, in a range of $10\ \Omega$ to $13\ \Omega$ settling time and inductor current almost remains same. When load regulation is found out for this circuit it is found to be 29.82 %.

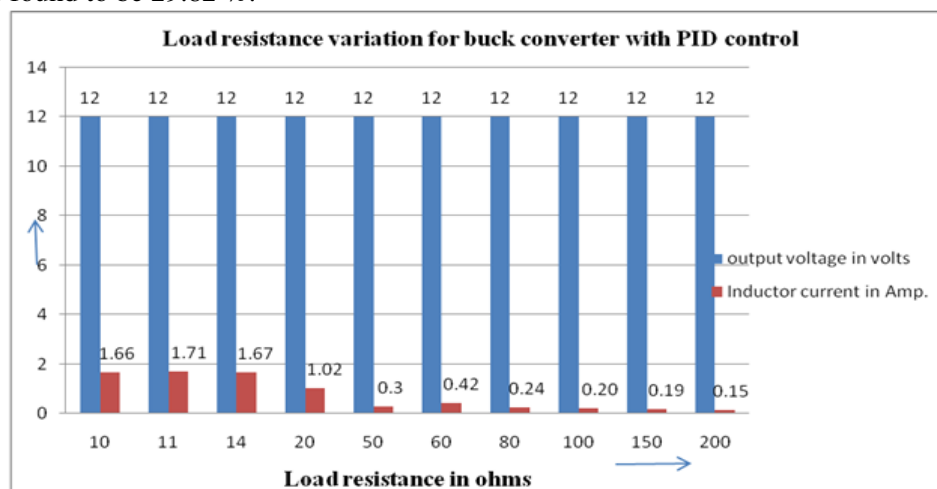


Figure No.3.4.1.1 Bar graph for variation of load resistance in PID control circuit.

3.6 Effect of variation of line voltage on buck converter with PID control

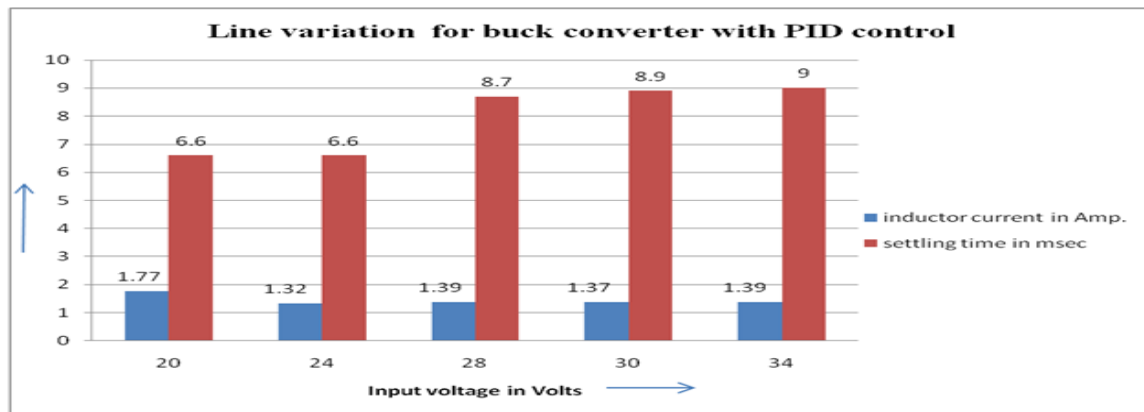


Figure No. 3.4.1.2 Bar graph line variation.

Figure 3.4.1.2 shows the line variation for PID controlled buck converter circuit. As input voltage increases, inductor current also increases and settling time also increases. The settling time is in msec for this circuit. From this variation we can say that this controller can be used in range of 20 V to 28 V with same output voltage, settling time and inductor current.

IV. SMC CONTROL METHOD

From above all control methods sliding mode control is the only non linear method and its performance is studied for comparison with other linear methods. SMC could be implemented for switch mode power supplies. The controller block diagram of SMC is shown in figure no. 4.1

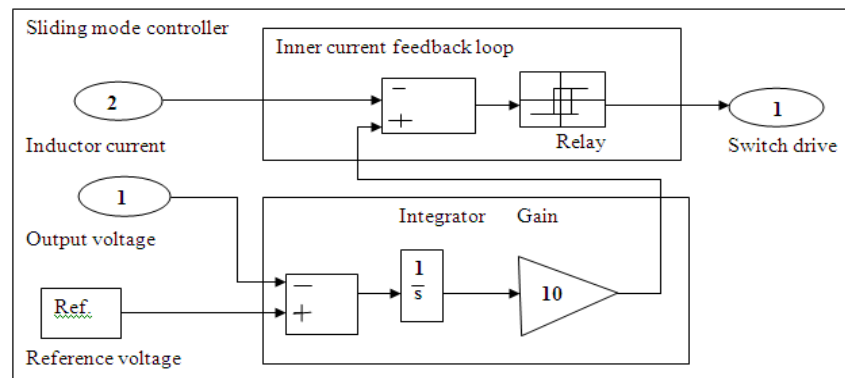


Figure No. 4.1 The simulation controller block diagram SMC.

4.1 Selection of various parameters for the circuit

The control topology consists of a linear and non-linear part. The non-linear parameter can be selected, while it is left to the designer to tune the linear part and get the optimum values depending on the application. The output of the integral is amplified through a gain and the result is subtracted from the inductor loop, the difference is passed through a hysteresis. One major drawback of this model is the lack of a standard procedure to select the gain. The hysteresis parameter can be selected by measuring the peak-to-peak inductor current and these are the limits for the hysteresis parameters.

Table 2:- Main Circuit Parameters:-

Parameter name	symbol	Value
Input voltage	V_{in}	24 volts
Output voltage	V_o	12 volts
Capacitor	C	220 μ F
Inductor	L	69 μ H
Load resistance	R_L	13 Ω
Nominal switching frequency	f	100 kHz
Switch_off	SW1	u= 0
Switch_on	SW1	u= 1

4.2. Buck converter with sliding mode control simulated circuit diagram

Considering above circuit in which a buck converter when considered with SMC controller it has been observed that the circuit has settling time of 20 msec. The output voltage attends steady state value of 12 V, which is expected output from this application. Under the load variation of SMC circuit from 0 to ∞ , it was found that as load resistance increases load current decreases and settling time increases continuously

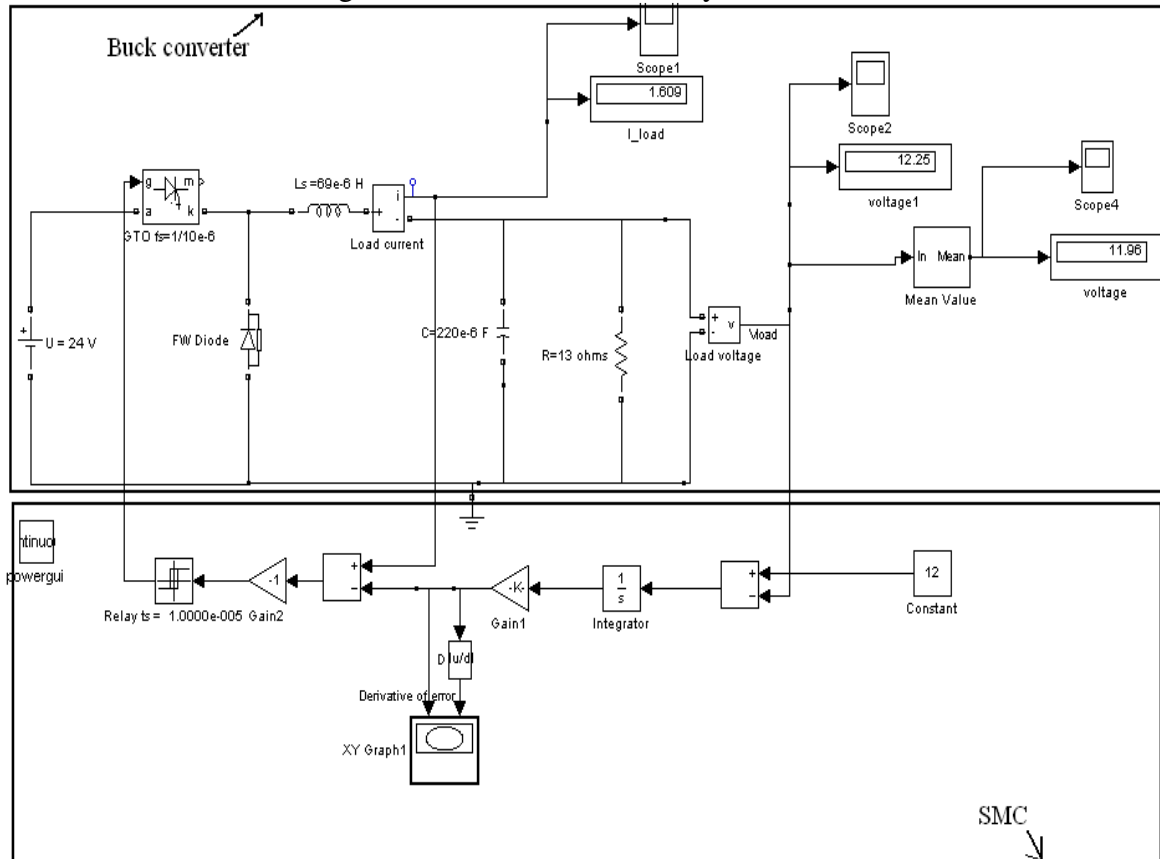


Figure No. 4.2.1 Simulation diagram for buck converter with SMC

4.2.1 Effect of variation of load resistance on buck converter with SMC control

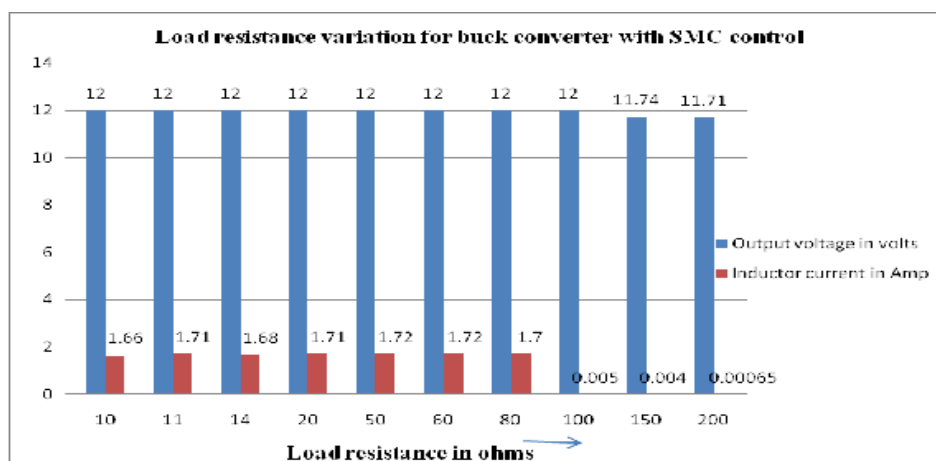


Figure No.4.2.1.1 Bar graph for load resistance variation

Above bar graph shows the effect of load variation on buck converter with SMC controller. As resistance value increases inductor current decreases. For ∞ resistance voltage is 23.69 and inductor

current is 1.16×10^{-10} A. But in the range of 10Ω to 13Ω load resistances inductor current and load voltage almost remain constant.

4.2.2 Effect of variation of line voltage on buck converter with PID control

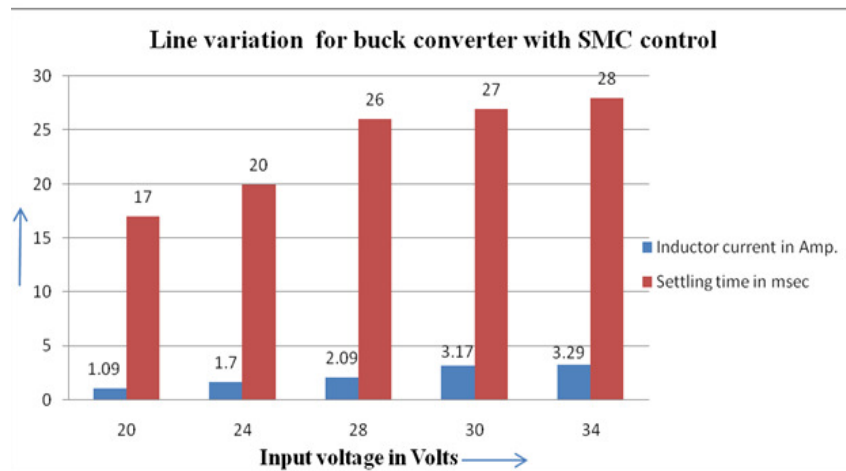


Figure No.4.2.2.1 Bar graph for line variation

V. MATLAB SIMULATION MODEL OF BUCK CONVERTER WITH SMC PID CONTROLLER

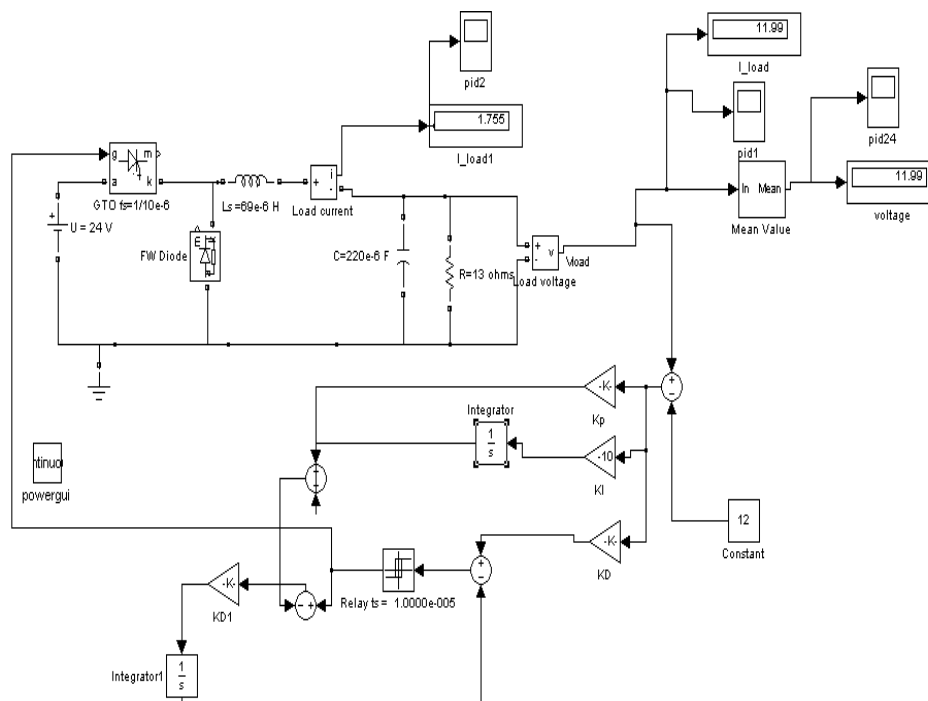


Figure No.5.1 Buck converter with SMC PID controller.

Above figure shows simulated model of buck converter with SMC PID controller. In this model SMC and PID controllers both are considered to get advantages of both control methods. From performance comparison of SMC PID with other controllers we can say that this circuit has large settling time but very less overshoot or no overshoot in voltage. Whenever we can consider this settling time and required more accuracy we can go for SMC PID model.

5.1 Effect of variation of load resistance on buck converter with SMC PID control:-

Above bar graph shows the effect of load variation on buck converter with SMC PID controller. As resistance value increases inductor current decreases. For ∞ resistance voltage is 23.69 V and inductor current is 1.06e-12 A. But in the range of 10 Ω to 15 Ω load resistances inductor current and load voltage almost remain constant

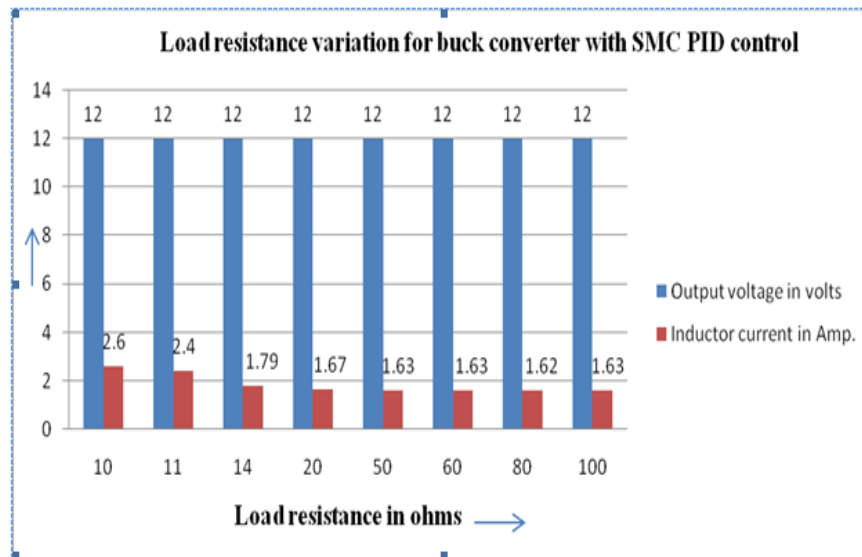


Figure No.5.1.1 Bar graph for load resistance variation

5.2 Effect of variation of line voltage on buck converter with SMC PID control

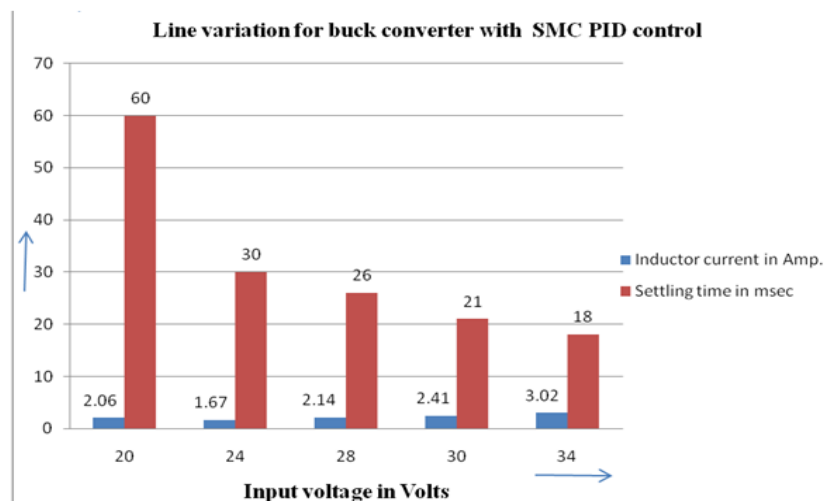


Figure No.5.2.1 Bar graph for line variation

Performance comparison

Table1 shows the summary of the performance characteristics of the buck converter between PI, PID, SMC and SMC PID controller quantitatively. Based on the data tabulated in Table 71, PID has the fastest settling time of 2.5 msec while SMC has the slowest settling time of 20 m seconds. An extra of 17.5 m seconds is required for the SMC controller for steady state voltage.

Table 1

Sr.No	Type of circuit	Output voltage			Inductor current		
		Rise Time	Maximum Peak overshoot	Setting Time	Rise Time	Maximum Peak overshoot	Setting Time
1	Buck Converter	0.9 m sec	11V	4 m sec	0.9 m sec	2A	4 m sec
2	Buck Converter with PI controller	0.25 msec	25 V	6.5 msec.	0.25 msec	40 A	6.5 msec.
3	Buck Converter with PID controller	0.2m sec	16V	2.5 m sec	0.1 m sec	28 A	2 m sec
4	Buck Converter with SMC controller	3 m sec	0.5 V	20 m sec	3 m sec	3.5 A	20 m sec
5	Buck Converter with SMC PID controller	3 m sec	0.1V	30m sec	2 m sec	3.2 A	30 m sec

VI. COMPARISON GRAPH FOR RISE TIME, DELAY TIME AND SETTLING TIME FOR ALL EXISTING CONTROLLERS

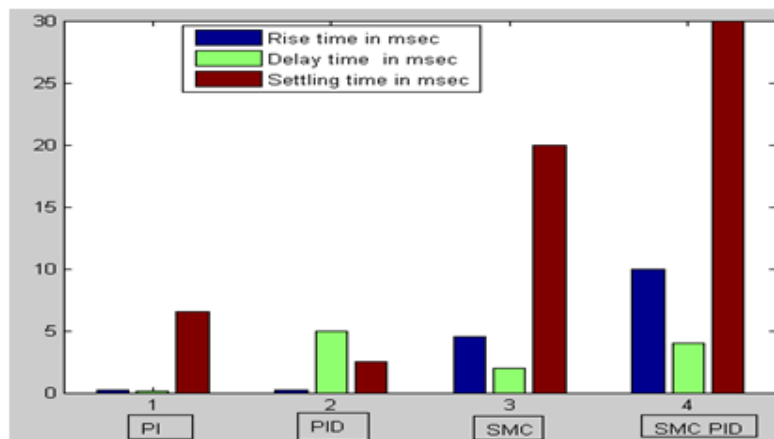


Figure No. 6.1.comparative graph for all existing controllers

6.1 Comparative graph for peak overshoot, regulation, output voltage and inductor current all existing controllers

From comparison we can say for same output voltage and inductor current peak overshoot is maximum for PI control and no overshoot for SMC PID control method. From the performance analysis of uncompensated buck converter we can say that because of disturbances and nonlinearities output voltage of converter is 14.12 V instead of 12 V.

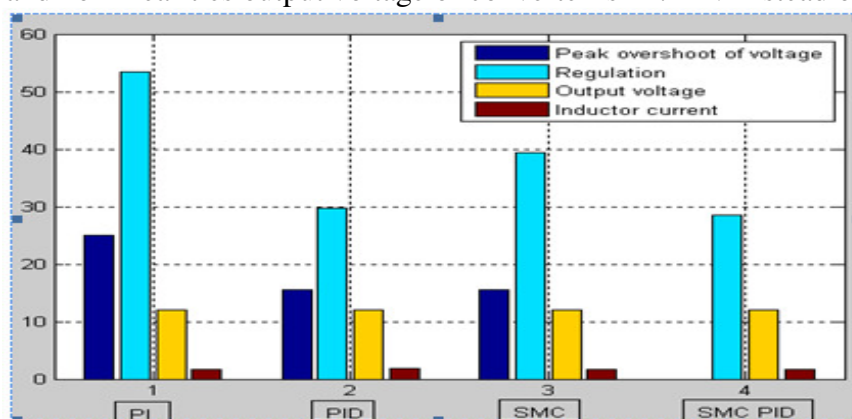


Figure No.6.1.1. Comparative graph for all existing controllers.

VII. CONCLUSION

As SMC is not operating at a constant switching frequency and converters have a highly nonlinear and time varying nature therefore it is selected to control such kind of DC- DC converter. Therefore it is also selected as control technique for performance analysis. The waveforms of simulated output

voltage and current were obtained, studied and compared with the waveforms from other controllers for performance comparison. By studied references papers in details the waveforms were found to be in precise proximity of theoretical waveforms. Some concluding points which are analyzed in following points. From performance comparison of SMC with PI and PID it was found that it has large settling time. So when more voltage accuracy is required and large settling time can be considered then we can go for SMC or SMC PID control method. But when less cost, less accuracy and less complexity is required, than PI or PID control method can be used. When buck converter is considered with PI control within 6.5 msec output voltage attends 12 V.

ACKNOWLEDGEMENT

We wish to acknowledge the support given by Principal, Modren College of Engineering, Pune for carrying out the present research work and HOD Mrs. Prof. N.R. Kulkarni department of Electrical Engg. for constant encouragement.

REFERENCES

- [1]. M.Ahmed, M.Kuisma, P. Silventoinen, "Implementing Simple Procedure for Controlling Switch Mode Power Supply Using Sliding Mode Control as a Control Technique", XIII-th International Symposium on Electrical Apparatus and technologies (Siela). May 2003, pp 9-14, Vol. 1
- [2]. Hongmei Li and Xiao Ye "Sliding-Mode PID Control of DC-DC Converter" , 5th IEEE Conference on Industrial Electronics and Applications .
- [3]. V.I.Utkin, Sliding modes and their application in variable structure systems, MIR Publishers, Moscow, 1978
- [4]. R. Venkataramanan, A. Sabanovic, S. Cuk:"Sliding-mode control of DC-to-DC converters," IECON Conf. Proc., 1985, pp. 251-258.
- [5]. G. Spiazzi, P. Mattavelli, L. Rossetto, L. Malesani, "Application of Sliding Mode Control to Switch-Mode Power Supplies," Journal of Circuits, Systems and Computers (JCSC), Vol. 5, No. 3, September 1995, pp.337-354.
- [6]. Siew-Chong Tan, Member, IEEE, Y. M. Lai, Member, IEEE, and Chi K. Tse, Fellow, IEEE "Indirect Sliding Mode Control of Power Converters .

Biography

Shelgaonkar(Bindu) Arti Kamalakar was born in Aurangabad, India, in Year 1978. She received the Bachelor in electrical engg. degree from the University of Dr. BAMU Aurangabad city, in Year 1999 and the pursing Master in 2008 degree from the University of Pune , in Year, both in control system engineering.



N.R. Kulkarni received the Bachelor in electrical engg. degree from WCE ,Sangli in 1985, M.E.(Electrical) Control System from COEP Pune in 1998, Ph.D.(Electrical) in 2011. Area of interest control System, Electrical machine, Nonconventional energy, Nonlinear system, Sliding mode control.



PERFORMANCE EVALUATION OF DS-CDMA SYSTEM USING MATLAB

Athar Ravish Khan

Department of Electronics & Telecommunication

Jawaharlal Darda Institute of Engineering and Technology, Yavatmal, Maharashtra, India

ABSTRACT

The author evaluated the performance of synchronous DS-CDMA systems over multipath fading channel and AWGN Channel. The synchronous DS-CDMA system is well known for eliminating the effects of multiple access interference (MAI) which limits the capacity and degrades the BER performance of the system. This paper investigated the bit error rate (BER) performance of a synchronous DS-CDMA system over AWGN and Rayleigh channel, which is affected by the different number of users, as well as different types spreading codes. The promising simulation results explore the comparative study of different DS-CDMA system parameter and showed the possibility of applying this system to the wideband channel. Different MATLAB functions and MATLAB program segments are explained for the simulation of DS-CDMA system.

KEYWORDS: CDMA system, QPSK, BER, Rayleigh Channel, AWGN channel, MATLAB program segment, Gold Sequence, M- sequence.

I. INTRODUCTION

Direct-sequence code-division multiple access (DS-CDMA) is currently the subject of much research as it is a promising multiple access capability for third and fourth generations mobile communication systems. Code-division multiple access (CDMA) is a technique whereby many users simultaneously access a communication channel. The users of the system are identified at the base station by their unique spreading code. The signal that is transmitted by any user consists of the user's data that modulates its spreading code, which in turn modulates a carrier. An example of such a modulation scheme is quadrature phase shift keying (QPSK). In this paper, we introduce the Rayleigh channel and AWGN Channel, and investigated the bit error rate (BER) performance of a synchronous DS-CDMA system over these channels. In the DS-CDMA system, the narrowband message signal is multiplied by a large bandwidth signal, which is called the spreading of a signal. The spreading signal is generated by convolving a M-sequence & GOLD sequence code with a chip waveform whose duration is much smaller than the symbol duration. All users in the system use the same carrier frequency and may transmit simultaneously. The receiver performs a correlation operation to detect the message addressed to a given user and the signals from other users appear as noise due to de-correlation. The synchronous DS-CDMA system is presented for eliminating the effects of multiple access interference (MAI) which limits the capacity and degrades the BER performance of the system. MAI refers to the interference between different direct sequences users. With increasing the number of users, the MAI grows to be significant and the DS-CDMA system will be interference limited. The spreading M & GOLD sequences in a DS-CDMA system need to have good cross-correlation characteristics as well as good autocorrelation characteristics [P. Alexander *et.al*],[E. Dinan *et.al*]. The goal is to reduce the fading effect by supplying the receiver with several replicas of the same

information signal transmitted over independently fading paths. The remainder of the paper is organized as follows. In the next section we present channel modelling. Section 3 deals with modulation and Demodulation scheme used in the system. Section 4 deals with proposed transmitter and receiver model for simulation. Different MATLAB functions, program segments and flow of program segment are explained in the Section 5 & 6 respectively, the paper ends with simulation results and conclusion.

II. CHANNEL MODEL

2.1. Rayleigh fading channel Model:

Rayleigh fading is a statistical model for the effect of a propagation environment on a radio signal, such as that used by wireless devices. Rayleigh fading models assume that the magnitude of a signal that has passed through such a transmission medium will vary randomly, or fade, according to a Rayleigh distribution the radial component of the sum of two uncorrelated Gaussian random variables. [C.Trabelsi et.al]. Rayleigh fading is viewed as a reasonable model for tropospheric and ionospheric signal propagation as well as the effect of heavily built-up urban environments on radio signals. Rayleigh fading is most applicable when there is no dominant propagation along a line of sight between the transmitter and receiver Rayleigh fading is a reasonable model when there are many objects in the environment that scatter the radio signal before it arrives at the receiver, if there is sufficiently much scatter, the channel impulse response will be well modelled as a Gaussian process irrespective of the distribution of the individual components. If there is no dominant component to the scatter, then such a process will have zero mean and phase evenly distributed between 0 and 2π radians. The envelope of the channel response will therefore be Rayleigh distributed [Theodore S. Rappaport].

2.2 . AWGN channel Model

Additive White Gaussian Noise channel model as the name indicate Gaussian noise get directly added with the signal and information signal get converted into the noise in this model scattering and fading of the information is not considered[Theodore S. Rappaport].

III. MODULATOR AND DEMODULATOR

A QPSK signal is generated by two BPSK signals. To distinguish the two signals, we use two orthogonal carrier signals. One is given by $\cos 2\pi f_c t$, and the other is given by $\sin 2\pi f_c t$. A channel in which $\cos 2\pi f_c t$ is used as a carrier signal is generally called an in-phase channel, or *Ich*, and a channel in which $\sin 2\pi f_c t$ is used as a carrier signal is generally called a quadrature-phase channel, or *Qch*. Therefore, $d_I(t)$ and $d_Q(t)$ are the data in *Ich* and *Qch*, respectively. Modulation schemes that use *Ich* and *Qch* are called quadrature modulation schemes. The mathematical analysis shows that QPSK [X.Wang et..al]

$$s_n(t) = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_c t + (2n-1)\frac{\pi}{4}\right) \quad \text{for } n = 1, 2, 3, 4 \quad (1)$$

This yields the four phases $\pi/4$, $3\pi/4$, $5\pi/4$ and $7\pi/4$ as needed. This results in a two-dimensional signal space with unit basis functions. The even Equation(2) and odd Equation(3) samples of signal are given by,

$$\phi_1(t) = \sqrt{\frac{2}{T_s}} \cos(2\pi f_c t) \quad (2)$$

The first basis function is used as the in-phase component of the signal and the second as the quadrature component of the signal. An illustration of the major components of the transmitter and receiver structure is shown below.

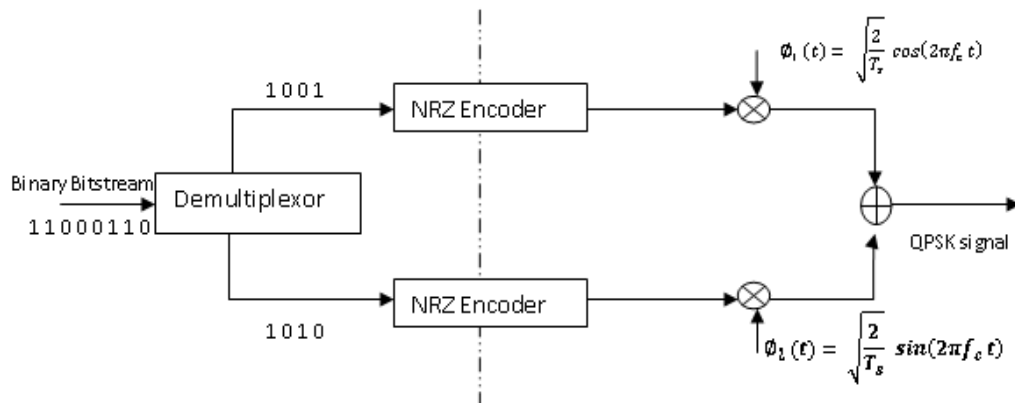


Figure.1 QPSK Modulator

The binary data stream is split into the in-phase and quadrature-phase components. These are then separately modulated onto two orthogonal basis functions. In this implementation, two sinusoids are used. Afterwards, the two signals are superimposed, and the resulting signal is the QPSK signal. Note the use of polar non-return-to-zero encoding. These encoders can be placed before for binary data source, but have been placed after to illustrate the conceptual difference between digital and analog signals involved with digital modulation. In the receiver structure for QPSK matched filters can be replaced with correlators. Each detection device uses a reference threshold value to determine whether a 1 or 0 is detected as shown in the Figure (2)

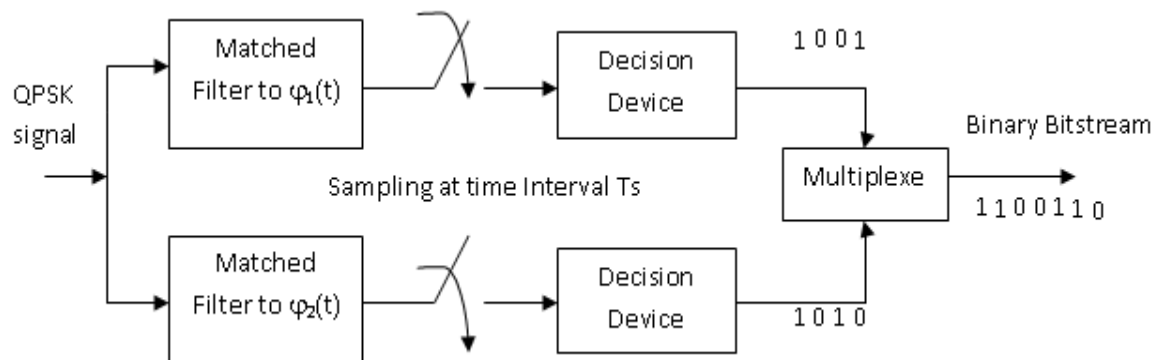


Figure.2 QPSK Demodulator

IV. PROPOSED SYSTEM MODEL

4.1 Proposed Transmitter Model:

The randomly generated data in system can be transmitted with the help of proposed transmitter model which is shown in Figure(3) given below

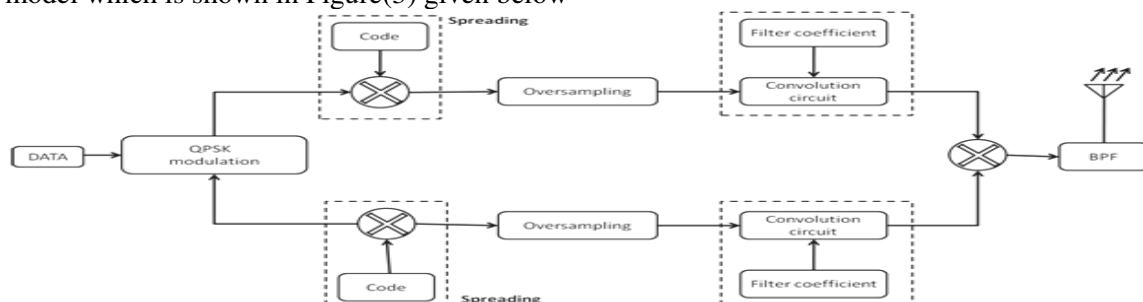


Figure.3 DS-CDMA transmitter

At first, the data generator generates the data randomly, that generated data is further given to the mapping circuit. Mapping circuit which is consisting of QPSK modulator converts this serially random data into two parallel data streams even and odd samples i.e. I_{ch} (in-phase) and Q_{ch} (quadrature phase) [X.Wang.*et.al*]. This I_{ch} and Q_{ch} are then convolved with codes and spreaded individually by using M-sequence or Gold sequence codes. The spreaded data is given to the oversampler circuit which converts unipolar data into bipolar one, then this oversampled data is convolved using with help of filter coefficients of T-filter. Then these two individual data streams are summed up and passed through Band pass filter (BPF) which is then transmitted to channel.

4.2 Proposed Receiver Model:

The randomly generated data in system which is transmitted through the channel can be received with the proposed receiver model which is shown in Figure (4) given below.

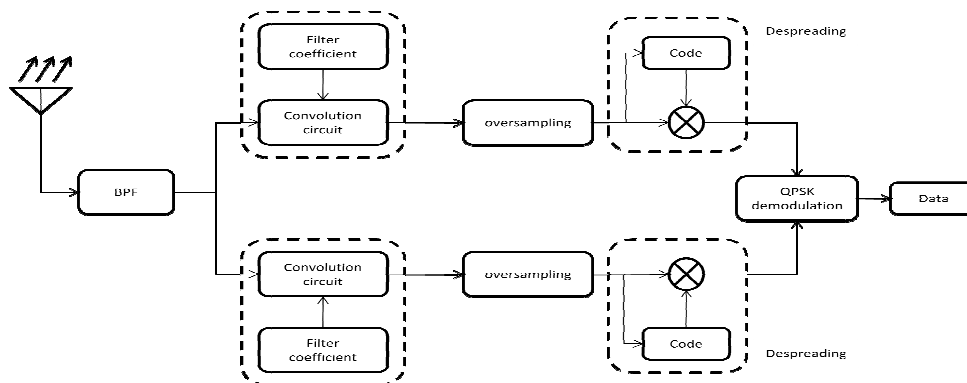


Figure.4 DS-SS receiver

At the receiver, the received signal passes through band pass filter (BPF), where spurious signals are eliminated. Then the signal is divided into two streams and convolved using filter coefficients, by which Inter Symbol Interference (ISI) in the signal is eliminated. This signal is despread using codes, also synchronized. These two despread streams are then fed to a Demapping circuit, which consists of a QPSK demodulator. The demodulator circuit converts the two parallel data streams into a single serial data stream. Thus, the received data is recovered at the end.

V. MATLAB SIMULATIONS

5.1 DS-SS System:

This section shows the procedure to obtain BER of a synchronous DS-SS. In the synchronous DS-SS, users employ their own sequences to spread the information data. At each user's terminal, the information data are modulated by the first modulation scheme. Then, the first bits of the modulated data are spread by a code sequence, such as an M-sequence or a Gold sequence. The spread data of all the users are transmitted to the base station at the same time. The base station detects the information data of each user by correlating the received signal with a code sequence allocated to each user. In the simulation, QPSK is used as the modulation scheme. The parameters used for the simulation are defined as follows [Hiroshi Harada *et.al*]:

```
sr = 2560000.0;    % symbol rate
ml = 2;           % number of modulation levels
br = sr * ml;      % bit rate
nd = 200;         % number of symbol
ebn0 = [0:20];    % Eb/No
irfn = 21;        % number of filter taps
IPOINT = 8;       % number of oversample
alfs = 0.5;       % roll off factor
```


The coefficient of the filter is defined as given in the above program segment ,evaluates the performance of QPSK and the MATLAB function *hrollfcoef* is use to evaluate the filter coefficient based on the above parameter.

```
[xh]=hrollfcoef(irfn,IPOINT,sr,alfs,1);
           % T Filter Function
[xh2]=hrollfcoef(irfn,IPOINT,sr,alfs,0 );
           % R Filter Function
```

The parameter for the spread sequences, namely M-sequence and Gold sequences are used. By denoting variables as *seq 1, or 2* a code sequence is selected. Next, the number of registers is set to generate an M-sequence. In synchronous DS-CDMA, the number of code sequences that can be allocated to different users is equal to the number of code lengths. Therefore, the length of the code sequence must be larger than the number of users. To generate a code sequence, we must specify the number of registers, the position of the feedback tap, and the initial values of the registers. To generate a Gold sequence and an orthogonal Gold sequence, two M-sequences are needed. Therefore, the following parameters are used. By using these parameters, a spread code is generated, and the generated code is stored as variable *code*.

```
user = 3           % number of users
seq = 1;           % 1:M-sequence 2:Gold
stage = 3;         % number of stages
ptap1 = [1 3];     % position of taps for 1st
ptap2 = [2 3];     % position of taps for 2nd
regi1 = [1 1 1];   % initial value of register for 1st
regi2 = [1 1 1];   % initial value of register for 2nd
```

Here, *code* is a matrix with a sequence of the number of users multiplied by the length of the code sequence. An M-sequence is generated by MATLAB function *mseq.m*, and a Gold sequence is generated by MATLAB function *goldseq.m*. An orthogonal Gold sequence can be generated by adding a 0 bit of data to the top or bottom of a Gold sequence. Because the generated code sequence consists of 0 and 1, the program converts it into a sequence consisting - 1 and 1.

```
switch seq
case 1           % M-sequence
    code = mseq(stage,ptap1,regi1,user);
case 2           % Gold sequence
    m1 = mseq(stage,ptap1,regi1);
    m2 = mseq(stage,ptap2,regi2);
    code = goldseq(m1,m2,user);
end

code = code * 2 - 1;
clen = length(code);
```

When *rfade* is 0, the simulation evaluates the BER performance in an AGWN channel. When *rfade* is 1, the simulation evaluates the BER performance in a Rayleigh fading environment [C.Trabelsi et.al].

```
rfade = 1;         % Rayleigh fading 0:nothing 1:consider
itau = [0,8];      % delay time
dlv11 = [0.0,40.0]; % attenuation level
n0 = [6,7];        % number of waves to generate fading
th1 = [0.0,0.0];   % initial Phase of delayed wave
itnd1 = [3001,4004]; % set fading counter
now1 = 2;          % number of directwave + delayed wave
tstp = 1 / sr / IPOINT / clen; % time resolution
fd = 160;          % doppler frequency [Hz]
```

```
flat = 1; % flat Rayleigh environment
itndel = nd * IPOINT * clen * 30; % number of fading counter to skip
```

Then, the number of simulation loops is set. The variables that count the number of transmitted data bits and the number of errors are initialized.

```
nloop = 10; % simulation number of times
noe = 0;
nod = 0;
```

The transmitted data in the in-phase channel and quadrature phase modulated by QPSK are multiplied by the code sequence used to spread the transmitted data. The spread data are then oversampled and filtered by a roll-off filter and transmitted to a communication channel. Here, MATLAB functions *compoversamp2.m*, *compconv2.m* and *qpskmod.m* used for oversampling filtering, and modulation, filter parameter *xh* form T-filter is provided in *compconv2* function.

```
data = rand(user,nd*ml) > 0.5;
[ich, qch] = qpskmod(data,user,nd,ml); % QPSK modulation
[ich1,qch1] = spread(ich,qch,code); % spreading
[ich2,qch2] = compoversamp2(ich1,qch1,IPOINT); % over sampling
[ich3,qch3] = compconv2(ich2,qch2,xh); % filter
```

Above program segment demonstrate the transmitter section of the DS-CDMA system. During this process *ich1,qch1* get transformed into *ich3* and *qch3*. The signals transmitted from the users are synthesized by considering the *if-else* statement depending upon the number of user *ich4* and *qch4* is generated

```
if user == 1 % transmission based of Users
    ich4 = ich3;
    qch4 = qch3;
else
    ich4 = sum(ich3);
    qch4 = sum(qch3);
end
```

The synthesized signal is contaminated in a Rayleigh fading channel as shown in below program segment. In reality, the transmitted signals of all users are contaminated by distinctive Rayleigh fading. However, in this simulation, the synthesized signal is contaminated by Rayleigh fading. Function *sefade.m* used to consider the Rayleigh fading

```
if rfade == 0
    ich5 = ich4;
    qch5 = qch4;
else % fading channel
    [ich5,qch5] = sefade(ich4,qch4,itaу,dlv1,th1,n0,itnd1,now1,...
        ..length(ich4),tstp,fd,flat);
    itnd1 = itnd1 + itndel;
end
```

At the receiver, AWGN is added to the received data, as shown in the simulation for the QPSK transmission in Program Segment (5). Then, the contaminated signal is filtered by using a the root roll-off filter. Below program segment calculate the attenuation and add AWGN to the signal *ich6* and *qch6* and transform the signal to *ich8* and *qch8* using the filter coefficient *xh2*.

```
spow = sum(rot90(ich3.^2 + qch3.^2)) / nd; % attenuation Calculation
attn = sqrt(0.5 * spow * sr / br * 10^(-ebn0(i)/10));
snrlnr=10.^(ebn0(i)/10);
attnNEW=sum(attn)/400;
```

```

[ich6,qch6] = comb2(ich5,qch5,attn);           % Add White Gaussian Noise (AWGN)
[ich7,qch7] = compconv2(ich6,qch6,xh2);       % filter
sampl = irfn * IPOINT + 1;
ich8 = ich7(:,sampl:IPOINT:IPOINT*nd*clen+sampl-1);
qch8 = qch7(:,sampl:IPOINT:IPOINT*nd*clen+sampl-1);

```

The resampled data are now the synthesized data of all the users. By correlating the synthesized data with the spread code used at the transmitter, the transmitted data of all the users are detected. The correlation is performed by Program,

```

[ich9 qch9] = despread(ich8,qch8,code);       % despreading

```

The correlated data are demodulated by QPSK. [Fumiyuki ADACHI] Then, the total number of errors for all the users is calculated. Finally, the BER is calculated.

```

demodata = qpskdemod(ich9,qch9,user,nd,ml);   % QPSK demodulation
noe2 = sum(sum(abs(data-demodata)));
nod2 = user * nd * ml;
noe = noe + noe2;
nod = nod + nod2;

```

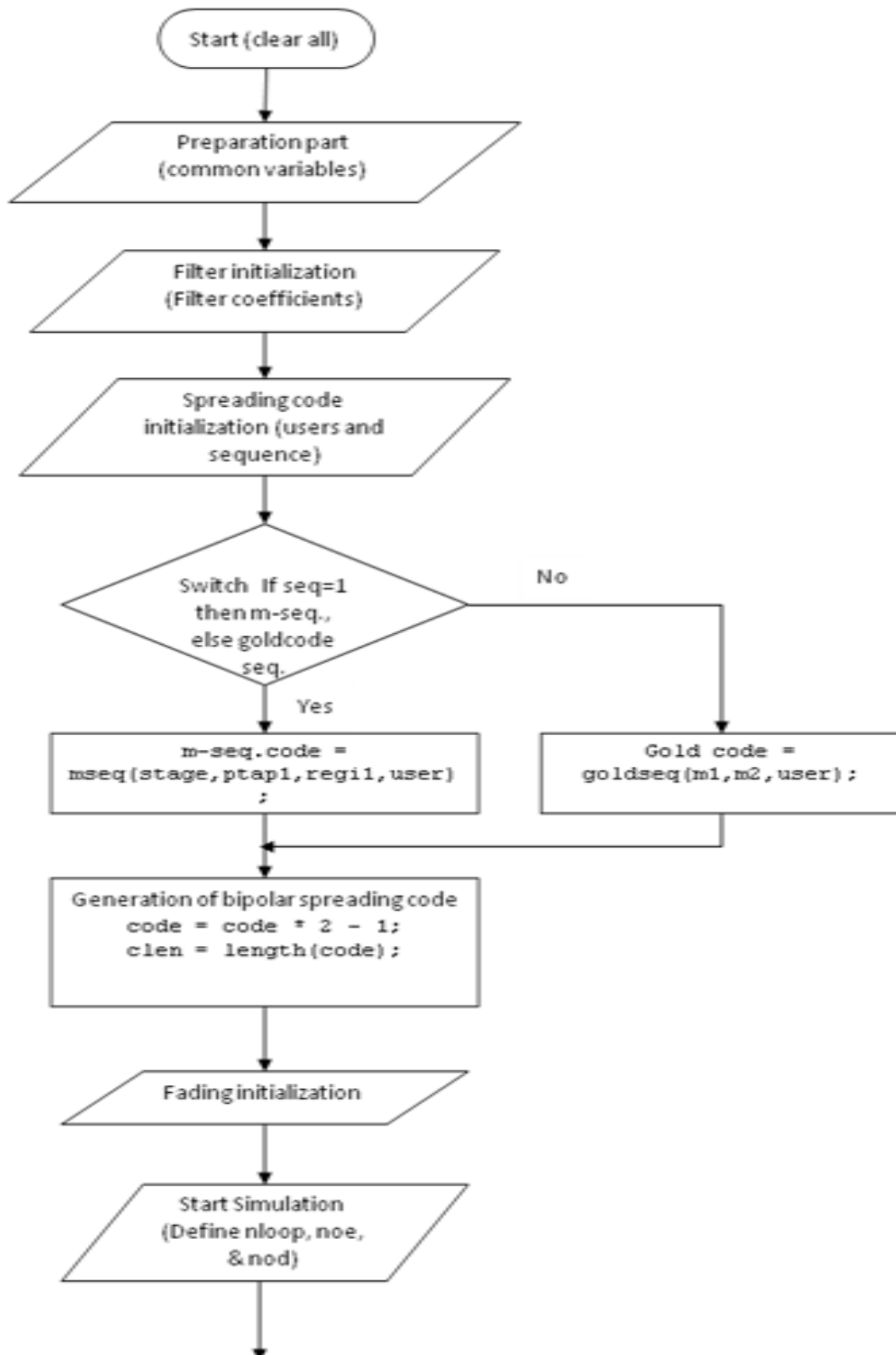
VI. SIMULATION FLOWCHART

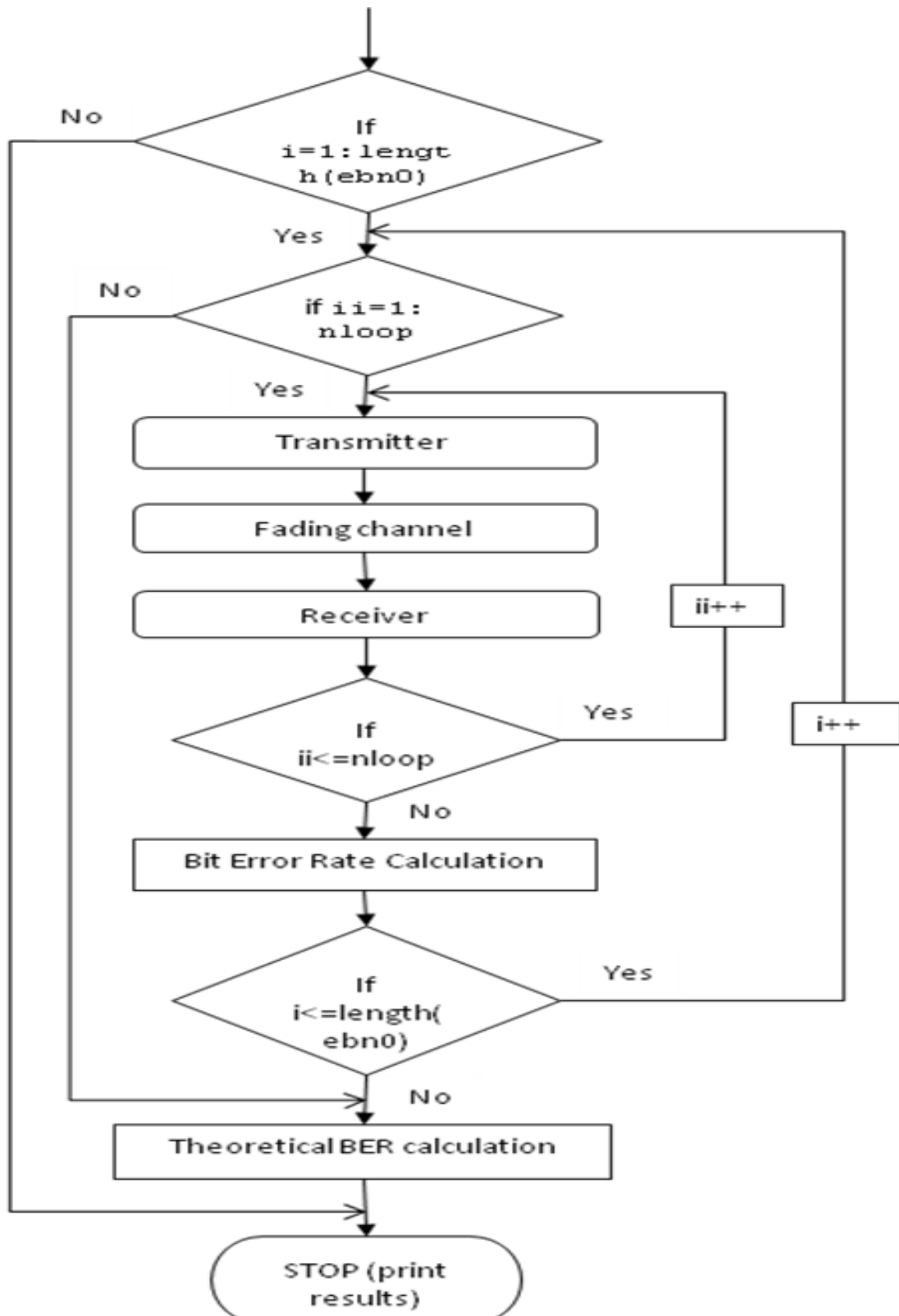
In order to simulate the system following step are

- Initialized the common variable
- Initialized the filter coefficient
- Select the switch for m-sequence and gold sequence
- Generate the spreading codes
- Initialize the fading by using variable rfade
- Define the variables for signal to noise ratio and the number of simulation requires as the data is random BER must have the average value of number of simulation.
- Simulate the system by using the proposed transmitter and receiver for different type of channel and codes
- Theoretical value of BER for Rayleigh channel and AWGN channel can be calculated by

$$BER_{theoretical}(AWGN) = \frac{1}{2} \operatorname{erfc}(\sqrt{E_b/N_0}) \quad \text{-----(3)}$$

$$BER_{theoretical}(Rayleigh) = \frac{1}{2} \left[1 - \frac{1}{\sqrt{1 + \frac{E_b}{N_0}}} \right] \quad \text{-----(4)}$$





VII. SIMULATION RESULTS OBTAINED

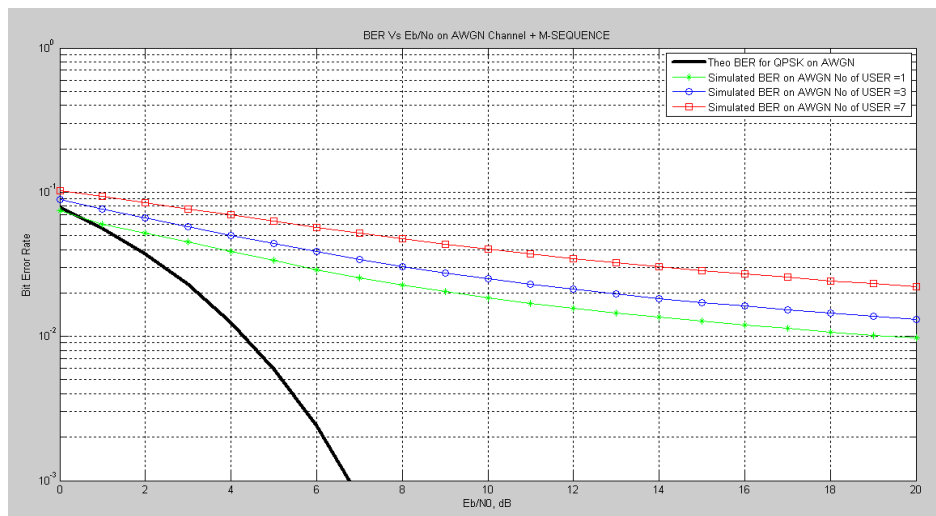


Figure.6 Performance of DS CDMA System in AWGN Environment With M Sequence

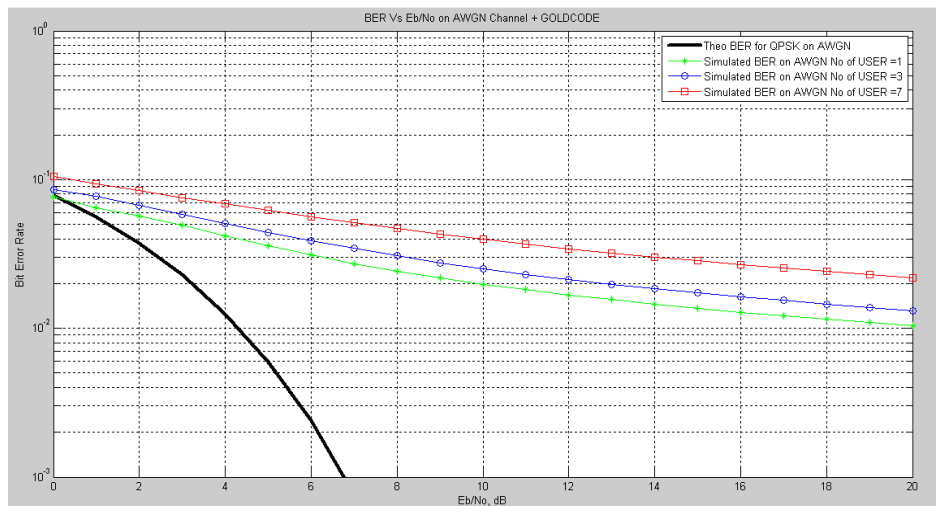


Figure.7 Performance of DS CDMA System in AWGN Environment With GOLD Sequence

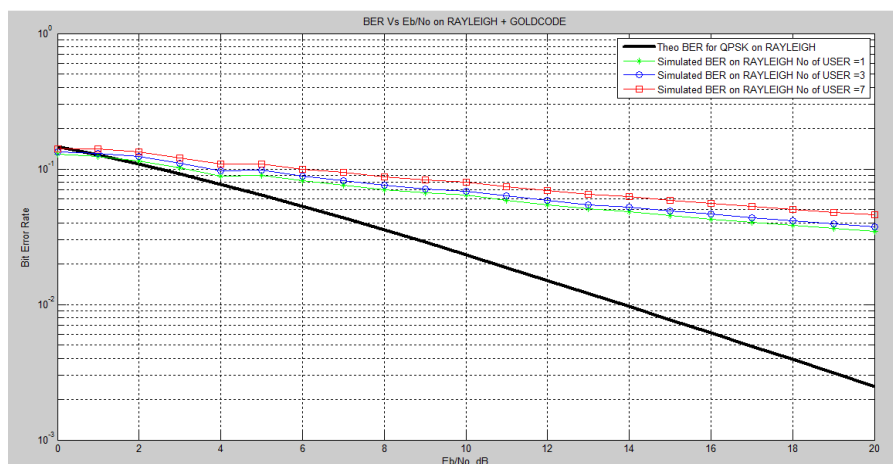


Figure.8 Performance of DS CDMA System in Rayleigh Environment With Gold Sequence

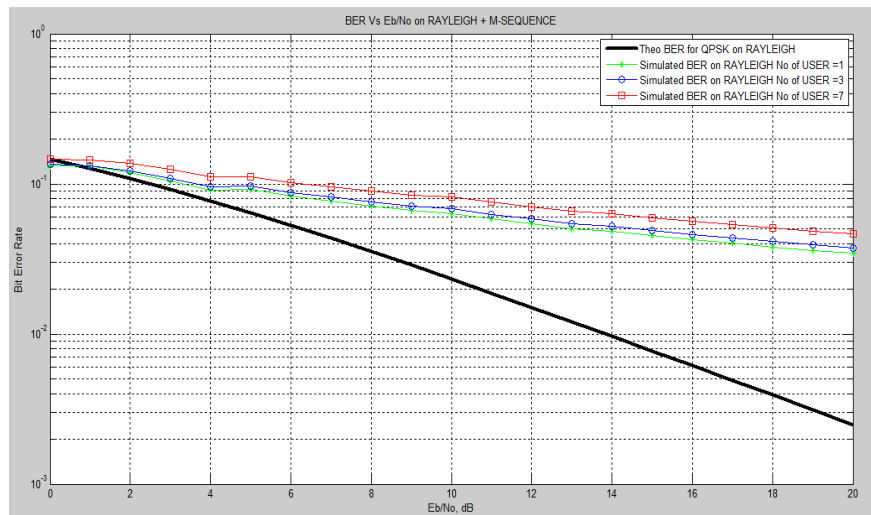


Figure.9 Performance of DS CDMA System in Rayleigh Environment With M Sequence

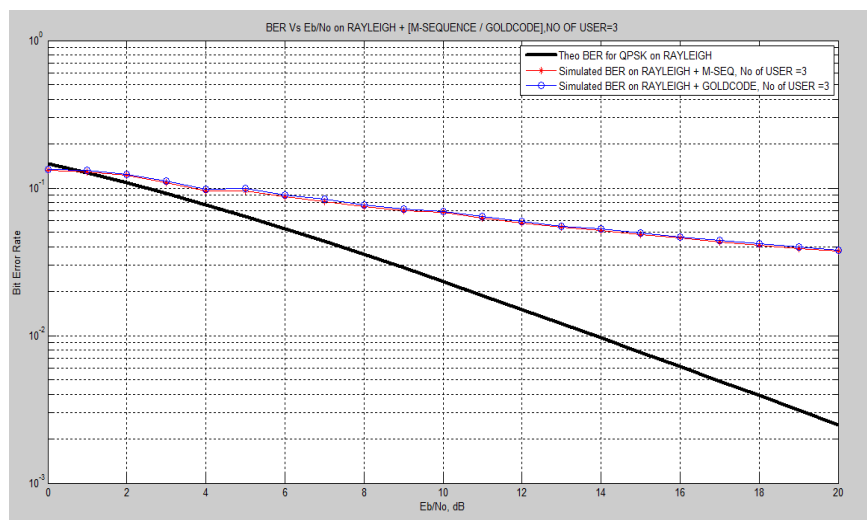


Figure.10 Performance of DS CDMA System in Rayleigh Environment With M & Gold Sequence

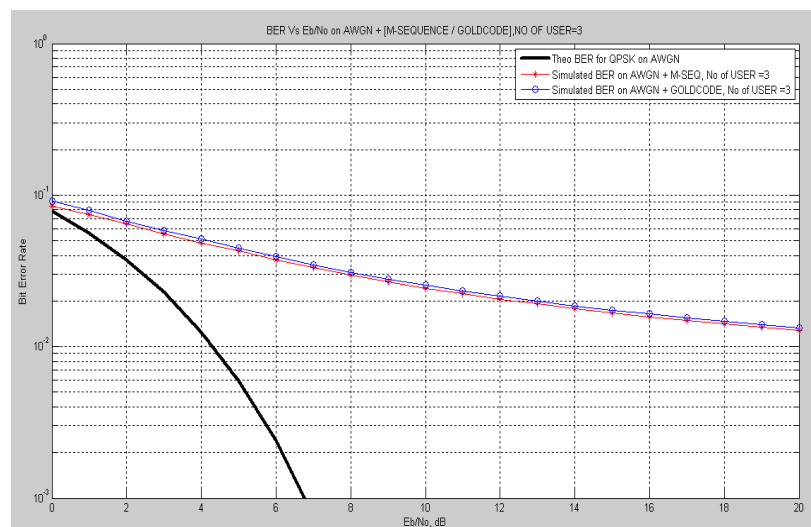


Figure.11 Performance of DS CDMA System in AWGN Environment With M & GOLD Sequence

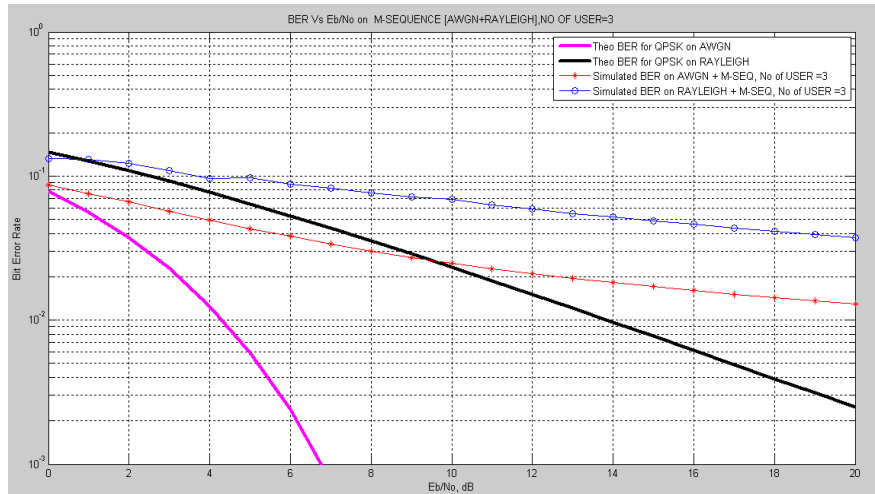


Figure.12 Performance of DS CDMA System in AWGN & Rayleigh Environment With M Sequence

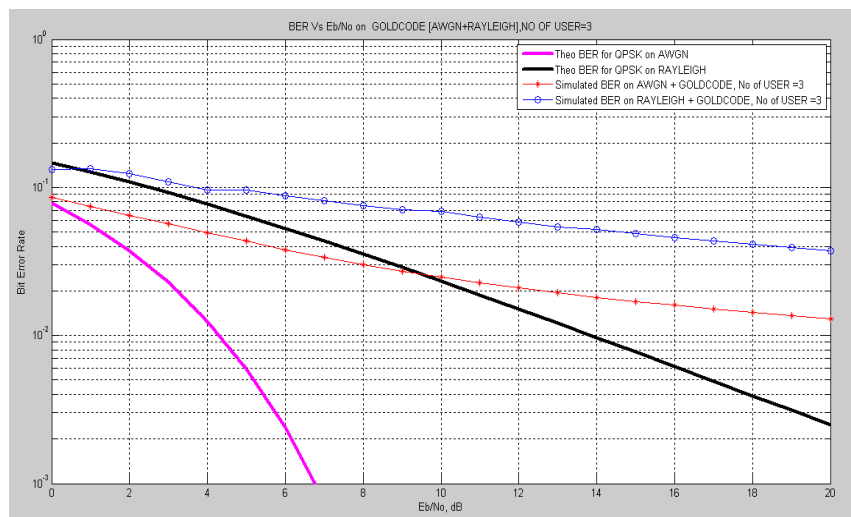


Figure.13 Performance of DS CDMA System In AWGN & Rayleigh Environment With Gold sequence

VIII. RESULTS AND CONCLUSION

In AWGN environment, when gold sequence or m sequence is used, for the different users the practical BER value for the minimum number of user is nearly approaches to the theoretical value of BER. In RAYLEIGH environment, when gold or m sequence is used, at the initial SNR value the practical and theoretical value of BER are same, as the SNR increases the practical BER value increases as compared to the theoretical value of BER. When the m sequence and gold sequence is considered in RAYLEIGH environment, at initial state the practical BER value and theoretical BER is same. But as the SNR increases, the practical BER value increases rapidly as compared to the theoretical BER value. When the m sequence and gold sequence is considered in AWGN environment, with single user, initially the practical BER value is same as the theoretical value, and with increasing SNR the practical value increases as compared to the theoretical value of BER. When either sequence is used in the system for AWGN and Rayleigh environment, initially the BER theoretical and practical value are nearly same. But, as the SNR value increases in case of the AWGN, the practical BER value increases rapidly as compared to the theoretical value, and in case of Rayleigh the practical value approaches to the theoretical value.

ACKNOWLEDGMENTS

The authors would like to thank firstly, our GOD, and all friends who gave us any help related to this work. Finally, the most thank is to our families and to our country INDIA which born us.

REFERENCES

- [1] Dr. Mike Fitton, Mike Fitton, "*Principles of Digital Modulation Telecommunications*" Research Lab Toshiba Research Europe Limited.
- [2] P. Alexander, A. Grant and M. C. Reed, "*Iterative Detection Of Code-Division Multiple Access With Error Control Coding*" European Trans.
- [3] Hiroshi Harada and Ramjee Prasad, "*Simulation and Software Radio*" for Mobile Communication.
- [4] X.Wang and H.V.Poor, "*Wireless Communication Systems: Advanced Techniques for Signal Reception*".
- [5] J. Proakis, *Digital Communications*, McGraw-Hill, McGraw-Hill.
- [6] Sklar B., "*A Structured Overview Of Digital Communications - A Tutorial Review - Part I*", IEEE Communications Magazine, August 2003.
- [7] Sklar B., "*A Structured Overview Of Digital Communications - A Tutorial Review - Part II*", IEEE Communications Magazine, October 2003.
- [8] E. Dinan and B. Jabbari, "*Spreading Codes for Direct Sequence CDMA and Wideband CDMA Cellular Networks*", IEEE Communications Magazine.
- [9] Shimon Moshavi, Bellcore, "*Multi-user Detection for DS-CDMA Communications*", IEEE Communications Magazine.
- [10] Hamed D. Al-Sharari, "*Performance of Wideband Mobile Channel on Synchronous DS-CDMA*", College of Engineering, Aljof University Sakaka, Aljof, P.O. Box 2014, Kingdom Of Saudi Arabia.
- [11] Theodore S. Rappaport, "*Wireless Communications Principles And Practice*".
- [12] Wang Xiaoying "*Study Spread Spectrum In Matlab*" School of Electrical & Electronic Engineering Nanyang Technological University Nanyang Drive, Singapore 639798.
- [13] Zoran Zvonar and David Brady, "*On Multiuser Detection In Synchronous CDMA Flat Rayleigh Fading Channels*" Department of Electrical and Computer Engineering Northeastern University Boston, MA 02115.
- [14] C.Trabelsi and A. Yongacoglu "Bit-error-rate performance for asynchronous DS-CDMA over multipath fading channels" *IEE Proc.-Commun.*, Vol. 142, No. 5, October 1995
- [15] Fumiyuki ADACHI "Bit Error Rate Analysis of DS-CDMA with joint frequency –Domain Equalization and Antenna Diversity Combinning" *IEICE TRANS.COMMUN.*, VOL.E87-B ,NO.10 OCTOBER 2004

Athar Ravish Khan was born in Maharashtra, INDIA, in 1979. He received the B.E. degree in electronics and telecommunication engineering, M.E. degree in digital electronics from SGBA University Amravati Maharashtra India, in 2000 and 2011 respectively. In 2000, he joined B.N College of Engineering Pusad and worked as lecturer. In 2006 he joined as lecturer in J.D Institute of Engineering and Technology Yavatmal, Maharashtra INDIA and in March 2011 he became an honorary Assistant Professor there. He is pursuing Ph.D. degree under the supervision of Prof. Dr. Sanjay M. Gulhane. His current research interests include digital signal processing, neural networks and wireless communications, with specific emphasis on UWB in underground Mines-Tunnels.



RECENT PHILOSOPHIES OF AGC OF A HYDRO-THERMAL SYSTEM IN DEREGULATED ENVIRONMENT

L. ShanmukhaRao¹, N.Venkata Ramana²,

¹Associate Professor, E.E.E.Department, Dhanekula Institute of Engineering & Technology,
Ganguru,Vijayawada, AP, India.

²Professor, E.E.E Department, JNTU Jagityal, AP, India.

ABSTRACT

In restructured power system, engineering aspects of planning and operation have to be reformulated although essentials ideas remain same. With emergency of distinct identities of GENCO's, TRANSCO's, DISCO's, and the ISO's many of the ancillary services of the vertically integrated utility will have a different role to play and hence have to be modified differently. Among these, ancillary services are, "the automatic generation control (AGC)".An attempt is made in this paper to present critical literature review and an up-to-date and exhaustive bibliography on the AGC of a hydro thermal system in deregulated environment. Various control aspects concerning the AGC problem have been highlighted.

KEYWORDS: Automatic generation control, Hydro-Thermal system, Deregulation.

I. INTRODUCTION

In modern power system network there are number of generating utilities interconnected together through tie-lines. In order to achieve integrated operation of a power system, an electric energy system must be maintained at a desired operating level characterized by nominal frequency, voltage profile and load flow configuration.

Modern power system normally consists of a number of subsystems interconnected through tie lines. For each subsystem the requirements usually include matching system generation to system load and regulating system frequency [5]. This is basically known as load-frequency control problem or automatic generation control (AGC) problem. It is desirable to achieve a better frequency constancy than is obtained by speed governing system alone. In an interconnected power system, it is also desirable to maintain the tie line flow at a given level irrespective of the load change in any area. To accomplish this, it becomes necessary to manipulate the operation of main steam valves or hydro gates in accordance with a suitable control strategy, which in turn controls the real power output of the generators. The control of the real power output of electric generators in this way is termed as "Automatic Generation Control (AGC)".This paper discuss the critical literature review of AGC schemes of hydro thermal system in deregulated environment.

II. AUTOMATIC GENERATION CONTROL

Power system loads are sensitive to frequency and following system frequency changes the aggregate load change follows deviation. When a generating unit is tripped or additional load is added to the system, the power mismatch is initially compressed by an extraction of the kinetic energy from the system inertial storage that causes a system frequency drop. [19]

As the frequency decreases, the power consumed by loads also decreases. Equilibrium for large system can be obtained when the frequency sensitive reduction of loads balances the power output of

tripped unit or that delivered to the additional load resulting in new frequency. This effect could stop the frequency decline in less than a couple of seconds. However, if the mismatch causes the frequency to deviate beyond the governor dead band of the generating units their outputs will be increased by the governor action. For such mismatches equilibrium is obtained when reduction in power consumed by the loads plus the increased generation due to governor action compensate the mismatch. Such equilibrium is normally obtained with a dozen seconds of the frequency incident. Governor droop is the percentage change in frequency that would cause unit generation to change by 100% of its capability. Typically speed droops for active generators are in the range of about 4%. With this level of frequency sensitivity and at the expense of some frequency deviation, generation adjust by governors provide ample opportunity for follow up manual control of units.

This automatic adjustment of generation by free governor action is known as primary frequency regulation. The objectives of the follow up control especially under normal changes of load are to return frequency to, schedule, to minimize production cost, and to operate the system at an adequate level of security.

Automatic generation control is a closed loop control system that particularly replaces this manual control. This form of generation control has become essential to the real time operation and control of interconnected power systems and operates in widely varying power system control environments ranging from autonomous to strongly interconnected systems with hierarchy multi-level control.

The purpose of AGC is to replace portions of the manual control.

As it automatically responds to normal load changes, AGC reduces the response time to a minute or more or less. Mainly due to delays associated with physically limited response rates of energy conversion, further reduction in the response of AGC is neither possible nor desired.

Neither follow up manual control nor AGC is able or expected to play any role in limiting the magnitude of their just frequency swing, which occurs within seconds after the loss of a block generation or load in the systems. For conditions where change of generation due to governor action and change of load due to its sensitivity to frequency are not enough to intercept the runaway frequency. Over and under frequency relay are among the last resorts for shedding loads to prevent system collapse or tripping generating units to prevent their damage.

The main aims behind the design of AGC are:

- a) The steady state frequency error following a step load perturbation should be zero.
- b) The steady state change in the tie flow following a step load change in an area must be zero.
- c) An automatic generation controller providing a slow monotonic type of generation responses should be preferred in order to reduce wear and tear of the equipment.

The objectives of AGC may, therefore be summarized as follows:

1. Each area regulates its own load fluctuations.
2. Each area assists the other areas, which cannot control their own load fluctuations.
3. Each area contributes to the control of the system frequency, so that the Operating costs are minimized.
4. The deviations in frequency and tie line power flow error to zero in the steady state.
5. When load changes are small, the system must be permitted to come back to the steady state (by natural damping) so that the mechanical power does not change small disturbances for economic reasons.

The problem of AGC can be subdivided into fast primary control and slow secondary control modes. The fast primary control (governing mechanism) mode tries to minimize the frequency deviations and has a time constant of the order of seconds. But, primary control does not guarantee the zero steady state error. The slow secondary control channel (supplementary control), with time constants of the order of minutes, regulates the generation to satisfy certain loading requirements and contractual tie-line loading agreements. The overall performance of the AGC in any power system depends on the proper design of both primary and secondary control loops.

The traditional power system industry has a “vertically integrated utility” (VIU) [9],[10],[14],[15],[32] structure and treated as a single utility company which monopolies generation, transmission and distribution in a certain geographic region. Interconnection between networks and

interaction between companies is usually voluntary to improve system reliability and performance. In the restructured or deregulated environment, vertically integrated utilities no longer exist. The first step in deregulation has been to separate the generation of power from the transmission and distribution, thus putting all the generation on the same footing as independent power producers (IPPs). So in the new scenario the utilities no longer own generation, transmission, and distribution; instead, there are three different entities, viz., GENCOs (generation companies), TRANSCOs (transmission companies) and DISCOs (distribution companies)

III. DEREGULATION

Analysis of the electrical industry beings with the reorganization of three components; Generation, Transmission and Distribution.

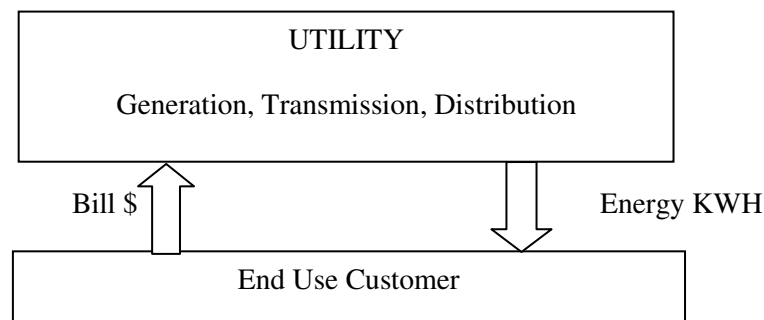


Fig.1 Schematic Diagram of Power System

Once electricity is generated,[3],[7] whether by burning fossil fuels, harnessing wind, solar or hydro energy, or through nuclear fission it is sent through high voltage, high capacity transmission lines to the local regions in which the electricity will be consumed. When electricity arrives in the region in which it is to be consumed, it is transformed to a lower voltage and sent through local distribution wires to end-use consumers. In general, all three of these vertically related sectors have typically been tied together within a utility, which has been either investor-owned or state de-regulated or owned by the municipality. For many years each sector was thought of as a natural monopoly.

Electric deregulation also known as electric restructuring, is the process by which the traditional monopoly structure for generating and delivering power to retail consumer is opened to competition by legislative or regulatory initiative. Addressed at the state level, electricity deregulation in its early stages and already beginning to benefits for consumers, the economy and future reliability of energy sources.

In the transmission and distribution sectors, effective competition would require that rival firms duplicate one another's wire network, which would be inefficient. If wires owned by different companies were allowed to interconnect to form a single network, the flow on one line affects the capacity of other lines in the system to carry power. The commodity that is opened to competition is called electricity generation or supply. Competitive suppliers offer it for sale. Customer can choose their competitive supplier. A customer's electricity bill will show a generation charge that represents the fee for the use of certain amount of electricity. Other elements the bills include amounts owned to the utility (now known as Distribution Company) for delivering the power to consumers through poles and wires. This delivery function is not being to competition.

Deregulation presents a chance to do better job at keeping cost down and making sure consumers have the kind of choice that best suits their needs. Historical deregulation or restructuring has the potential to produce gains in three broad sectors of the electricity utility industry operations, investment and consumption.

Concordia and Kirchmayer [1]-[3] have analyzed the AGC problem of two equal area thermal, hydro and hydro-thermal systems. They have studied the AGC of a hydro-thermal system considering non-reheat turbine and mechanical governor in hydro system, neglecting generation rate constraints. Their conclusion from simulation studies show that for minimum interaction between control areas frequency bias (B) must be set equal to area frequency response characteristics β . Although they have extensively studied the effect of variation of several parameters on dynamic performance of the

system, no explicit method has been suggested by them for the optimization of controllers. Concordia [3] has given the basic concepts of load-frequency control of an interconnected power system. He has discussed the effect of frequency bias and governor turbine model of a thermal system.

Cohn [7]-[8] has discussed mainly regarding the selection of frequency bias setting for large multi-area power system. His investigations reveal that for minimum interaction between control areas ideally the frequency bias setting (B) of a control area should match the combined generation and load frequency response i.e. area frequency response characteristics (β) of the area. However, Cohn has not addressed to the problem of optimum gain setting and structures of the supplementary controllers from the point of view of dynamic behavior of the system.

Nanda and Kaul [9],[10] have extensively studied the AGC problem of a two area reheat thermal system, using both parameter plane and ISE techniques for optimization of integral gain setting and for investigating the degree of stability of the system. They have studied the effect of GRC, area capacity effect, speed regulation parameter on optimum controller setting and system dynamic responses. The effect of variation of significant parameters on optimum controller setting and cost function has been brought out through sensitivity analysis neglecting GRC. However, they have not addressed to the problem pertaining to correction of time error and inadvertent interchange accumulations.

IEEE committee report on "Power Plant Responses" [13] shows that in practice GRC for reheat thermal system varies between 2.5% to 12% per minute and the permissible rate of generation for hydro plant is relatively much higher (a typical value of generation rate constraints (GRC) being 270% per minute for raising generation and 360% per minute for lowering generation), as compared to that for reheat type thermal units having GRC of the order of 3% per minute. Ref. [13] provides the transfer function model for steam and hydro turbines for AGC.

Nanda [14], [15], [32] have investigated the AGC problem of an interconnected hydro-thermal system in both continuous and discrete mode, with and without GRC. They are possibly the first to consider GRC to investigate the AGC problem of a hydro-thermal system with conventional integral controllers. They have found out the optimum integral controller settings and their sensitivity to GRC, speed regulation parameter 'R', base load condition etc. They have also studied the AGC problem of hydro thermal system, considering GRC where their main contribution is to explore the best value of speed regulation parameter. They have considered mechanical governor for hydro turbine.

V. Donde and M A. Pai, I.A. Hiskens [5] present AGC of a two area non reheat thermal system in deregulated power system. The concept of DISCO participation matrix (DPM) and area participation factor (APF) to represent bilateral contracts are introduced. However, they have not dealt with reheat turbine, GRC and hydro-thermal system in their work.

Bekhouché [25] has compared load frequency control before and after deregulation. Before deregulation ancillary services, including AGC are provided by a single utility company called a control area that owns both transmission and generation systems. After deregulation, the power system structure has changed allowing specialized companies for generation, transmission, distribution and independent system operator.

Richard D. Christie and Anjan Bose [20] have dealt with LFC (Load Frequency Control) issues in deregulated power system. It identifies the technical issues associated with load frequency control and also identifies technical solutions such as standards and algorithms, needed for the operation in this new restructured power system.

Meliopoulos, Cokkinides and Bakirtzis [23] have given the concept that in a deregulated environment, independent generators and utility generators may or may not participate in the load frequency control of the system. For the purpose of evaluating the performance of such a system, a flexible method has been developed and implemented. They proposed a method in which they assumed that load frequency control is performed by ISO (Independent System Operator) based on parameters defined by participating generating units. The participating units comprise utility generators and independent power producers. The utilities define the units which will be under load-frequency control, while the independent power producers may or may not participate in the load frequency control. For all the units which participate in the load-frequency control, the generator owner defines (a) generation limits, (b) rate of change and (c) economic participation factor. This information is transmitted to the ISO. This scheme allows the utilities to economically dispatch their own system, while at the same time permit the ISO to control the interconnected system operation. In the paper it has been shown

that if the percentage of units participating in this control action is very small, system performance deteriorates to a point that is unacceptable. It is therefore recommended that minimum requirements be established, based on system.

J. Kumar, Kah-Hoe Ng and G. Sheble[21][22] have presented AGC simulator model for price based operation in a deregulated power system. They have suggested the modifications required in the conventional AGC to study the load following in price based market operations. A framework for price-based operation is developed to assist in understanding AGC operation in the new business environment. The modified AGC scheme includes the contract data and measurements, which are continuous, regular and quiescent and hence, greatly improves control signals to unit dispatch and controllers. The proposed simulator is generic enough to simulate all possible types of load following contracts (bilateral and poolco). The proposed scheme includes ACE as a part of control error signal and thus, also satisfies the NERC performance criteria. The new framework requires establishment of standards for the electronic communication of contract data as well as measurements. They have highlighted salient differences between the automatic generation control in a vertical integrated electric industry (conventional scenario) and a horizontal integrated electric industry (restructured scenario). However, they have not addressed the aspects pertaining to reheat turbine, GRC and hydro-thermal system.

Donde, Pai and Hiskens [24] present AGC of a two area non-reheat thermal system in deregulated power system. In a restructured power system, the engineering aspects of planning and operation have to be reformulated although essential ideas remain the same. With the emergence of the distinct identities of GENCOs, TRANSCOs, DISCOs, many of the ancillary services of a vertically integrated utility will have a different role to play and hence have to be modeled differently. Among these ancillary services is the automatic generation control (AGC). In the new scenario, a DISCO can contract individually with a GENCO for power and these transactions are done under the supervision of the ISO or the RTO. In this paper, the two area dynamic model is formulated. Specifically it had focused on the dynamics, trajectory sensitivities and parameter optimization. The concept of a DISCO participation matrix (DPM) is proposed which helps the visualization and implementation of the contracts. The information flow of the contracts is superimposed on the traditional AGC system and the simulations revealed some interesting patterns. The trajectory sensitivities are helpful in studying the effects of parameters as well as in optimization of the ACE parameters viz. tie line bias and frequency bias parameter. The concept of DISCO participation matrix (DPM) and area participation factor (APF) to represent bilateral contracts are introduced.

IV. CONCLUSION

Literature survey shows that most of the earlier work in the area of automatic generation control in deregulated power system pertains to interconnected thermal system and no attention has been devoted to hydro-thermal systems involving thermal and hydro subsystems of widely different characteristics. The paper presents a critical review of AGC of hydro thermal system in deregulated environment. It has paid particular attention to categorize various AGC strategies in the literature that highlights its salient features. The authors have made a sincere attempt to present the most comprehensive set of references for AGC. It is anticipated that this document will serve as a valuable resource for any worker of the future in this important area of research.

REFERENCES

- [1] C. Concordia, L. K. Kirchmayer, "Tie-Line Power & Frequency Control of Electric Power Systems", AIEE Trans., vol. 72, part III, 1953, pp. 562-572.
- [2] C. Concordia, L. K. Kirchmayer, "Tie-Line Power & Frequency Control of Electric Power Systems-Part II, AIEE Trans., vol. 73, part III-A, 1954, pp. 133-141.
- [3] L. K. Kirchmayer, "Economic Control of Interconnected Systems", John Wiley, New York, 1959.
- [4] O. I. Elgerd, C. E. Fosha, "Optimum Megawatt Frequency Control of Multi-area Electric Energy Systems", IEEE Trans. on Power Apparatus and Systems, vol. PAS-89, No.4, Apr. 1970, pp. 556-563.
- [5] C. E. Fosha, O. I. Elgerd, "The Megawatt Frequency Control problem: A New Approach via Optimal Control Theory", IEEE Trans. on Power Apparatus and Systems, vol. PAS-89, No.4, Apr. 1970, pp. 563-574.

- [6] Nathan Cohn, "Some Aspects of Tie-Line Bias Control on Interconnected Power Systems", AIEE Trans., vol. 75, Feb. 1957, pp. 1415-1436.
- [7] Nathan Cohn, "Control of Generation & Power Flow on an Interconnected Power Systems", John Wiley, New York, 2nd Edition, July 1971.
- [8] IEEE Committee Report, "IEEE Standard Definition of Terms for Automatic Generation Control of Electric Power Systems", IEEE Trans. Power Apparatus and Systems, vol. PAS-89, Jul. 1970, pp. 1358-1364.
- [9] J. Nanda, B. L. Kaul, "Automatic generation Control of an Interconnected Power System", IEE Proc., vol. 125, No.5, May 1978, pp. 385-391.
- [10] "J. Nanda, M. L. Kothari, P. S. Satsangi, "Automatic Generation Control of an Interconnected Hydro-thermal System in Continuous and Discrete modes considering Generation Rate Constraints", IEE Proc., vol. 130, pt. D, No.1, Jan. 1983, pp 17-27.
- [11] D.G. Ramey, J. W. Skooglund, "Detailed Hydro governor representation for System stability Studies", IEEE Trans. on Power Apparatus and Systems, vol. PAS-89, No. Jan. 1970, pp. 106-112.
- [12] Power plant responses," IEEE Committee report, IEEE Trans. Power Apparatus & Systems, vol. PAS-86, Mar. 1967, pp. 384-395.
- [13] IEEE Committee Report, "Dynamic Models for steam and Hydro Turbines in Power System Studies", IEEE Trans. Power Apparatus & systems, Nov./Dec. 1973, pp. 1904-1915.
- [14] M. L. Kothari, B. L. Kaul, J. Nanda, "Automatic Generation Control of Hydro-Thermal System", Journals of Institute of Engineers (India), pt. EL-2, vol. 61, Oct. 1980, pp. 85-91.
- [15] M. L. Kothari, J. Nanda, P. S. Satsangi, "Automatic Generation Control of Hydro-Thermal System considering Generation Rate Constraint", Journals of Institute of Engineers (India), pt. EL, vol. 63, June 1983, pp. 289-297.
- [16] M. Leum, "The Development and Field Experience of a Transistor Electric Governor for Hydro Turbines," IEEE Trans. Power Apparatus & Systems, vol. PAS-85, Apr. 1966, pp. 393-402.
- [17] F. R. Schleif, A. B. Wilbor, "The Co-ordination of Hydraulic Turbine Governors for Power System Operation," IEEE Trans. Power Apparatus and Systems, vol. PAS-85, No.7, Jul. 1966, pp. 750-758.
- [18] L. Hari, M. L. Kothari, J. Nanda, "Optimum Selection of Speed Regulation Parameter for Automatic Generation Control in Discrete Mode considering Generation Rates Constraint", IEEE Proc., vol. 138, No.5, Sept 1991, pp. 401-406.
- [19] P. Kundur, "Power System Stability & Control," McGraw-Hill, New York, 1994, pp. 418-448.
- [20] Richard D. Christie, Anjan Bose, "Load Frequency Control Issues in Power System Operations after Deregulation", IEEE Transactions on Power Systems, Vol.11, No.3, August 1996, pp 1191-1196.
- [21] J. Kumar, Kah-Hoe Ng and G. Sheble, "AGC Simulator for Price-Based Operation: Part I", IEEE Transactions on Power Systems, Vol.12, No.2, May 1997, pp527-532
- [22] J. Kumar, Kah-Hoe Ng and G. Sheble, "AGC Simulator for Price-Based Operation: Part II", IEEE Transactions on Power Systems, Vol.12, No.2, May 1997, pp 533-538.
- [23] A.P Sakis Meliopoulos, G.J.Cokkinides and A.G.Bakirtzis," Load-Frequency Control Service in a Deregulated Environment", Decision Support Systems 24(1999) 243-250.
- [24] V. Donde, M. A. Pai and I. A. Hiskens, "Simulation and Optimization in an AGC System after Deregulation", IEEE Transactions on Power Systems, Vol. 16, No.3, August 2001, pp 481-488.
- [25] Dr.N.Bekhouche,"Automatic Generation Control Before and after Deregulation" IEEE 2002 Page 321-323
- [26] H.L.Zeynelgil, A.Demiroren, N.S.Sengor "Application of ANN technique to AGC for multi area system", electric power and energy system, page 345-354 July-2001
- [27] A.Demiroren, H.L.Zeynelgil, N.S.Sengor, "Application of ANN technique to load frequency control for three area power system" IEEE PTC,sept-2001
- [28] A.Demiroren, E.Yesil "Automatic Generation Control with fuzzy logic controllers in the power system including SMES units", electric power and energy system, 26(2004) page 291-305
- [29] S.P.Ghoshal," Optimizations of PID gains by particle swarm optimizations in fuzzy based automatic generation control", electric power and energy system, April 2004 page 203-212
- [30] S.P.Ghoshal," Application of GA/GA-SA based fuzzy automatic generation control of a multi-area thermal generating system", Electric Power System Research 70 (2004) 115-127.
- [31] H. Shayeghi, H.A. Shayanfar and O.P. Malik," Robust decentralized neural networks based LFC in regulated power system" Electric Power Systems Research 19 Apr 2006.
- [32] Manoranjan Parida and J. Nandal" Automatic Generation Control of a Hydro-Thermal System in Deregulated Environment", Proceedings of the Eighth International Conference on Electrical Machines and Systems, vol 2, Septmeber 2005 , page 942-947.

Authors

L. ShanmukhaRao received the Bachelor in Electrical and Electronics Engineering degree from the kakatiya University Warangal, A.P, in 2006 and the Master in Electrical Power engineering degree from the JNTUH,Hyderabad, in 2006. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, JNTUH, Hyderabad. His research interests include are Power System Operation and Control. He is currently Associate Professor at Dhanekula Institute of Engineering &Technology, Ganguru, Vijayawada, AP. India.



N. Venkata Ramana has received M. Tech from, S.V. University, India in 1991 and Ph.D. in Electrical Engineering from Jawaharlal Nehru Technological University (J.N.T.U),Hyderabad, India in Jan' 2005. His main research interest includes Power System Modeling and Control. He authored 2 books on power systems and published 14 research papers in national and international journals and attended 10 international conferences. He is currently Professor at J.N.T.U. College of Engineering, Jagityal, Karimnagar District, A.P., India.



DYNAMIC ROUTING SCHEME IN ALL-OPTICAL NETWORK USING RESOURCE ADAPTIVE ROUTING SCHEME

S. Suryanarayana¹, K. Ravindra², K. Chennakesava Reddy³

¹ Dept. of ECE, CMR Institute of Technology, JNT University Hyderabad, India

² Dept. of ECE, Mallareddy Institute of Tech & Science, JNT University Hyderabad, India

³ Dept. of EEE, TKR College of Engg & Tech, JNT University Hyderabad, India

ABSTRACT

With the increasing demand for high data transfer rate, the communication is getting new developments. For progressive data transfer at high data rate services, the means of communication has now taken high offering bandwidth architecture such as optical networks. In optical networks the mode of communication is completely an optical medium and data are transferred from various nodes to reach to the destination via optical routers. Though these networks have high bandwidth compatibility they offer heavy traffic congestion due to non-linear traffics resulting in degraded quality services. In this paper we present an adaptive methodology towards developing routing scheme in optical network based on queue based mechanism at wavelength router for comparatively higher offering quality of services.

KEYWORD: All-optical network, resource adaptive routing, dynamic routing scheme, throughput, overhead

I. INTRODUCTION

From the past ten to twenty years the usage of internet services are drastically increasing year by year. So we have to develop our communication systems to meet the demand for the data transfer. So to incorporate the quality of service for high data rate services, optical networks are the upcoming solution. Optical wavelength-division-multiplexing networks provide large bandwidth and are promising networks for the future Internet. Wavelength routed WDM systems that utilize optical cross connect are capable of switching data in the optical domain. In such systems, end-to-end all-optical light paths can be established and no optical-to electronic and electronic-to-optical conversions are necessary at intermediate nodes. Such networks are referred to as all-optical networks. Wavelength routed networks without wavelength conversion are also known as wavelength-selective (WS) networks [11]. In such a network, a connection can only be established if the same wavelength is available on all links between the source and the destination. This is the wavelength-continuity constraint. Wavelength routed networks with wavelength conversion are also known as wavelength-interchangeable (WI) networks [11]. In such a system, each router is equipped with wavelength converters so that a light path can be setup with different wavelengths on different links along the path. To establish a light path in a WDM network, it is necessary to determine the route over which the light path should be established and the wavelength to be used on all the links along the route. This problem is called the routing and wavelength assignment (RWA) problem. Routing and wavelength assignment requires that no two light paths on a given link may share the same wavelength. In addition, in WS networks, light paths must satisfy the wavelength continuity constraint, that is, the same wavelength must be used on all the links along the path. The RWA problem can be classified into two types: the static RWA problem and the dynamic RWA Problem. In the static RWA problem, the set of connections is known in advance, the problem is to set up light paths for the connections while minimizing network resources such as the number of wavelengths and

the number of fibers. Alternatively, one may attempt to set up as many light paths as possible for a given number of wavelengths.

Dynamic RWA tries to perform routing and wavelength assignment for connections that arrive dynamically. The objective of dynamic RWA is to minimize the blocking probability. The routing and wavelength assignment problem has been studied extensively.

A summary of the research in this area can be found in [16]. This problem is typically partitioned into two sub-problems: the routing sub-problem and the wavelength selection sub-problem [2, 3, 5, 6, 7, 10, 14]. For the routing sub-problem, there are basically three approaches, fixed routing [6], alternate routing [6, 7, 14], and dynamic adaptive routing [9, 10]. Among the routing schemes, dynamic adaptive routing, which is studied in this paper, offers the best performance. A large number of wavelength selection schemes have been proposed: random-fit [5], first-fit [5], least-used[15], most-used[15], min-product[8], least-loaded [11], max-sum[2, 15], and relative capacity loss[17].

The schemes can roughly be classified to three types. The first type, including random-fit and least-used, tries to balance the load among different wavelengths. The schemes in this category usually perform poorly in comparison to other types of RWA schemes. The second type, including first-fit, most-used, min-product, and least-loaded, tries to pack the wavelength usage. These schemes are simple and effective when the network state information is precise. The third type, including max-sum and relative capacity loss, considers the RWA problem from a global point of view. These schemes deliver better performance and are more computational intensive than the other types of schemes. In this study, we investigate the impact of route overhead information on the performance of the routing wavelength algorithms.

II. DYNAMIC ROUTING SCHEME

In this paper we outline the approach of providing quality of service based on route queue mechanism for higher quality of service. It is observed that the congestion probabilities at the link points are very heavy and a large computation is carried out at each router to provide an optimal routing. As the overhead in the route is basically due to packet blockage and queuing it is prime requirement to reduce this overhead to achieve high quality services. To achieve this objective in this paper we propose a markovian approach for a distributed optical network.

A queuing system consists of one or more routers that provide service of some sort to arriving node. Node who arrives to find all routers busy generally join one or more queues (lines) in front of the routers, hence the name queuing systems. There are several everyday examples that can be described as queuing systems [7], such as bank-teller service, computer systems, manufacturing systems, maintenance systems, communications systems and so on. Components of a Queuing System: A queuing system is characterized by three components: Arrival process - Service mechanism - Queue discipline.

2.1. Arrival Process

Arrivals may originate from one or several sources referred to as the calling population. The calling population can be limited or 'unlimited'. An example of a limited calling population may be that of a fixed number of machines that fail randomly. The arrival process consists of describing how node arrives to the system. If A_i is the inter-arrival time between the arrivals of the $(i-1)^{th}$ and i^{th} node, we shall denote the mean (or expected) inter-arrival time by $E(A)$ and call it

$$(\lambda) = 1/(E(A)) \text{ the arrival frequency.}$$

2.2. Service Mechanism

The service mechanism of a queuing system is specified by the number of routers (denoted by s), each server having its own queue or a common queue and the probability distribution of customer's service time. Let S_i be the service time of the i^{th} customer, we shall denote the mean service time of a customer by $E(S)$ and $\mu = 1/(E(S))$ the service rate of a server.

2.3. Queue Discipline

Discipline of a queuing system means the rule that a server uses to choose the next customer from the queue (if any) when the server completes the service of the current customer. Commonly used queue disciplines are:

FIFO - Node are served on a first-in first-out basis. LIFO - Node are served in a last-in first-out manner. Priority - Node are served in order of their importance on the basis of their service requirements.

2.4. Measures of Performance for Queuing Systems:

There are many possible measures of performance for queuing systems. Only some of these will be discussed here.

Let, D_i be the delay in queue of the i th customer W_i be the waiting time in the system of the i th customer = $D_i + S_i$ $Q(t)$ be the number of node in queue at time t $L(t)$ be the number of node in the system at time $t = Q(t) + \text{No. of node being served at } t$

Then the measures,

$$d = \lim_{L \rightarrow \infty} \frac{\sum_{L=1}^{L=2} D_L}{n}$$

$$w = \lim_{L \rightarrow \infty} \frac{\sum_{L=1}^{L=2} W_L}{n} \quad \text{----- (1)}$$

(if they exist) are called the steady state average delay and the steady state average waiting time in the system. Similarly, the measures,

$$Q = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T Q(t) dt$$

$$L = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T L(t) dt \quad \text{----- (2)}$$

(if they exist) are called the steady state time average number in queue and the steady state time average number in the system. Among the most general and useful results of a queuing system are the conservation equations:

$$Q = (\lambda) d \text{ and } L = (\lambda) w \quad \text{----- (3)}$$

These equations hold for every queuing system for which d and w exist. Another equation of considerable practical value is given by,

$$w = d + E(S) \quad \text{----- (4)}$$

Other performance measures are:

the probability that any delay will occur. - the probability that the total delay will be greater than some pre-determined value - that probability that all service facilities will be idle. - the expected idle time of the total facility. - the probability of turn-always, due to insufficient waiting accommodation.

2.5. Notation for Queues.

Since all queues are characterized by arrival, service and queue and its discipline, the queue system is usually described in shorten form by using these characteristics. The general notation is:

$$[A/B/s]:\{d/e/f\}$$

Where,

A = Probability distribution of the arrivals

B = Probability distribution of the departures

s = Number of routers (channels)

d = The capacity of the queue(s)

e = The size of the calling population

f = Queue ranking rule (Ordering of the queue)

There are some special notation that has been developed for various probability distributions describing the arrivals and departures. Some examples are,

M = Arrival or departure distribution that is a Poisson process

E = Erlang distribution

G = General distribution

GI = General independent distribution

Thus for example, the $[M/M/1]:\{\infty/\infty/FCFS\}$ system is one where the arrivals and departures are a Poisson distribution with a single server, infinite queue length, calling population infinite and the queue discipline is FCFS. This is the simplest queue system that can be studied mathematically. This queue system is also simply referred to as the M/M/1 queue.

III. SYSTEM DESIGN

The common characteristic of all markovian systems is that all interesting distributions, namely the distribution of the interarrival times and the distribution of the service times are exponential distributions and thus exhibit the markov (memoryless) property. From this property we have two important conclusions:

1. State of the system can be summarized in a single variable, namely the number of node in the system. (If the service time distribution is not memoryless, this is not longer true, since not only the number of node in the system is needed, but also the remaining service time of the customer in service.)
2. Markovian systems can be directly mapped to a *continuous time markov chain* (CTMC) which can then be solved.

3.1. The M/M/1-Queue

The M/M/1 Queue has iid interarrival times, which are exponentially distributed with specified parameters and also iid service times with exponential distribution. The system has only a single server and uses the FIFO service discipline. The waiting line is of infinite size. It is easy to find the underlying markov chain. As the system state we use the number of node in the system. The M/M/1 system is a pure birth-death system, where at any point in time at most one event occurs, with an event either being the arrival of a new customer or the completion of a customer's service. What makes the M/M/1 system really simple is that the arrival rate and the service rate are not state-dependent.

Steady-State Probabilities:

We denote the steady state probability that the system is in state $k(k \in \mathbb{N})$ by p_k , which is defined by

$$p_k := \lim_{t \rightarrow \infty} P_k(t) \quad \text{----- (5)}$$

$P_k(t)$ Where $p_k(t)$ denotes the (time-dependent) probability that there are 'k' node in the system at time t. The steady state probability p_k does not dependent on t. We focus on a fixed state k and look at the *flows* into the state and out of the state. The state k can be reached from state k-1 and from state k+1 with the respective rates $\lambda P_{k-1}(t)$ (the system is with probability $P_{k-1}(t)$ in the state k-1 at time t and goes with the rate from the predecessor state k-1 to state k) and $\mu P_{k+1}(t)$ (the same from state k+1). The total flow into the state k is then simply $\lambda P_{k-1}(t) + \mu P_{k+1}(t)$. The State k is left with the rate $\lambda P_k(t)$ to the state k+1 and with the rate $\mu P_k(t)$ to the state k-1 (for k=0 there is only a flow coming from or going to state 1). The total flow out of that state is then given by $\lambda P_k(t) + \mu P_k(t)$. The total rate of change of the flow into state k is then given by the difference of the flow into that state and the flow out of that state:

$$\frac{dP_k(t)}{dt} = (\lambda P_{k-1}(t) + \mu P_{k+1}(t)) - (\lambda P_k(t) + \mu P_k(t)) \quad \text{----- (6)}$$

Furthermore, since the p_k are probabilities, the *normalization condition*

$$\sum_{k=0}^{\infty} p_k = 1 \quad \text{----- (7)}$$

3.2. M/M/m-Queue

The M/M/m-Queue ($m > 1$) has the same interarrival time and service time distributions as the M/M/1 queue, however, there are m routers in the system and the waiting line is infinitely long. As in the M/M/1 case a complete description of the system state is given by the number of node in the system (due to the memoryless property). The M/M/m system is also a pure birth-death system.

3.3. M/M/1/K-Queue

The M/M/1/K-Queue has exponential inter arrival time and service time distributions, each with the respective parameters λ and μ . The nodes are served in FIFO-Order; there is a single server but the system can only hold up to K node. If a new customer arrives and there are already K nodes in the

system the new customer is considered lost, i.e. it drops from the system and never comes back. This is often referred to as *blocking*. This behavior is necessary, since otherwise (e.g. when the customer is waiting outside until there is a free place) the arrival process will be no longer markovian. As in the M/M/1 case a complete description of the system state is given by the number of node in the system (due to the memoryless property). The M/M/1/K system is also a pure birth-death system. This system is better suited to approximate “real systems” (like e.g. routers) since buffer space is always finite.

IV. RESULT OBSERVATION

For the evaluation of the suggested approach a distributed optical network environment is been developed.

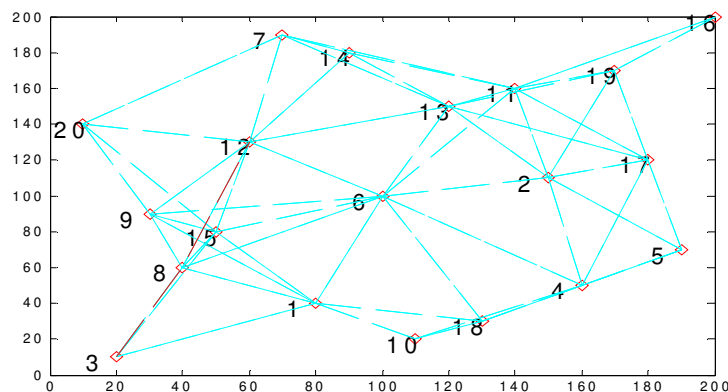


Fig 1: optical network architecture considered

The above figure illustrates about how the assigned nodes are established as a network. By applying the routing method we have got all the possible links in between the nodes. Whenever a node has to deliver packets to the destination from the source it has to follow the shortest path and reliable path to travel. This is done by the routing methodology. It is observed from the above figure that the reliable path is chosen and the data is passed through that path. It is represented by red dotted lines. After calculation of reliable path the data packets has to travel to destination in communication phase. The setup phase finds out which are all the possible paths shown in above figure.

The figure 2 plot is between number of data packets and data queue. It's been observed that delay performance is better in proposed system when compared to conventional systems. In proposed model we are using queuing methods to overcome the delay. It's clearly observed that as the number of data packets is increasing the queue management is good in the proposed work.

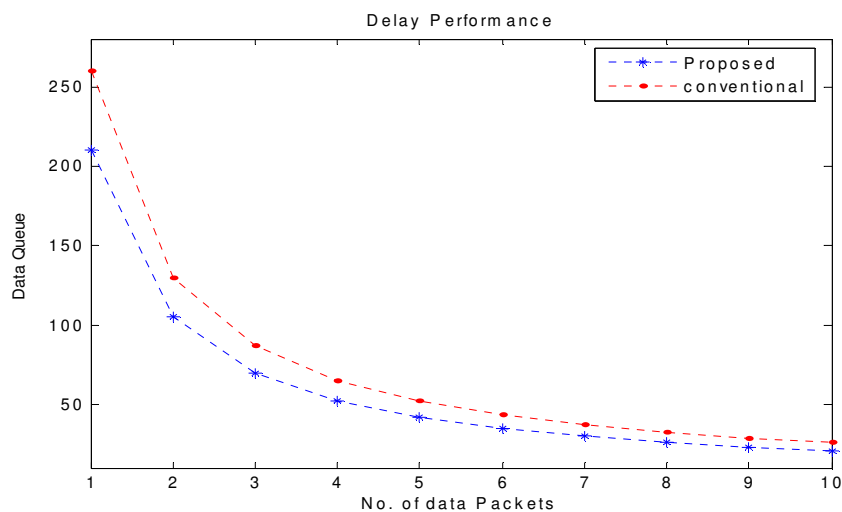


Fig 2: Delay Performance

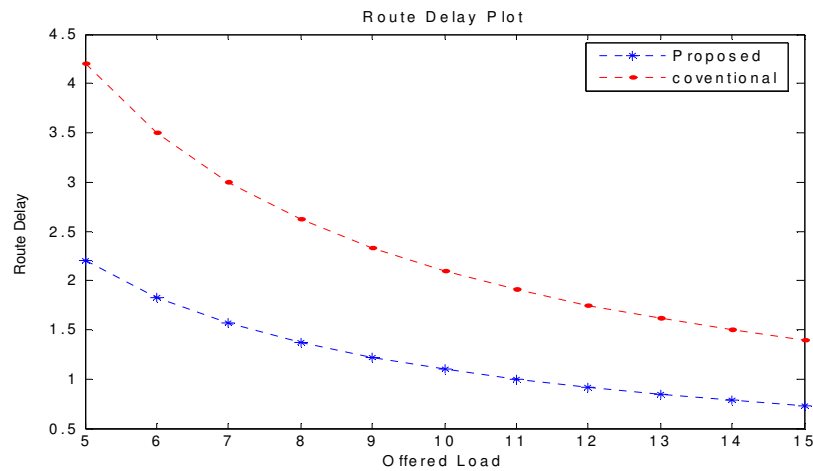


Fig 3: Route Delay Plot

Usually when the offered load is more the route delay will be there. The load is more means obviously the traffic and due to traffic congestion also will be more. In order to overcome the congestion in the network due to heavy traffic queuing models are used. The above plot mentions how the route delay varies when the offered load is increased. For the proposed method route delay is less when compared to the convention method.

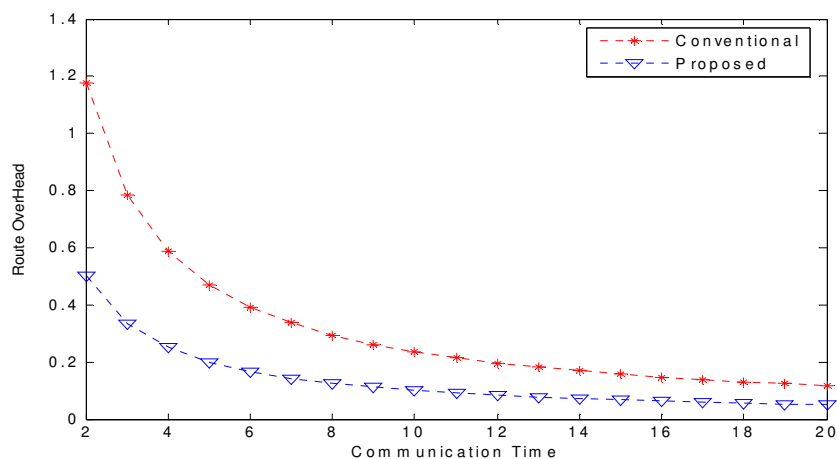


Fig 4: Route Overhead

Due to route delay the route overhead will increase. It leads to failure in data packets arrival. Chance of data packet loss will be there. Hence by applying queuing model the problem is clearly solved. It's observed that even increase in communication time the route overhead is less in proposed methodology when compared to conventional method.

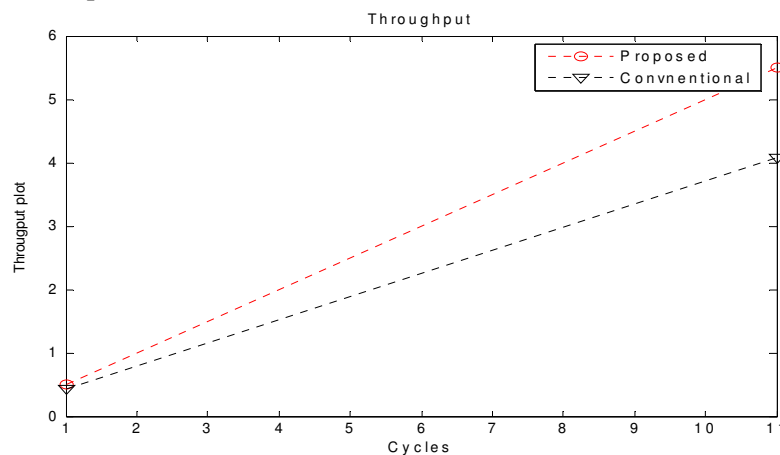


Fig 5: Throughput Plot

For any system throughput is the main parameter to be concentrated on. The above plot gives idea that the routing system which is used without any queuing model has got less throughput when compared to the reliable model which we have proposed. The throughput is comparatively high when compared to the conventional method. Under a similar observation for different topology of network is simulated and observed as,

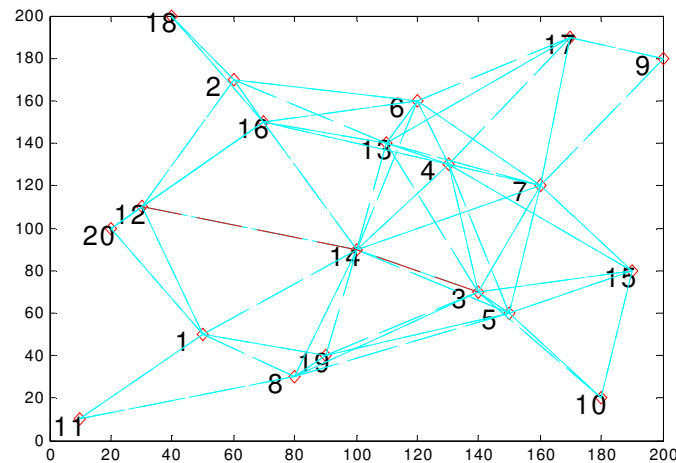


Fig 6: Link Probability

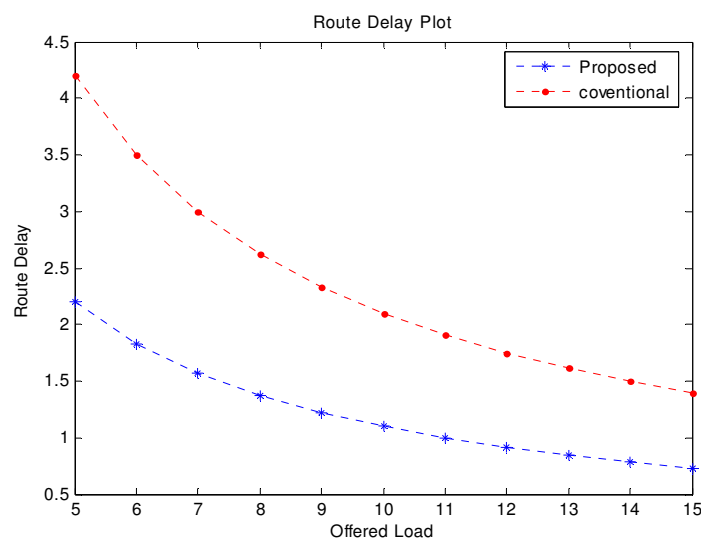


Fig 7: Route Delay Plot

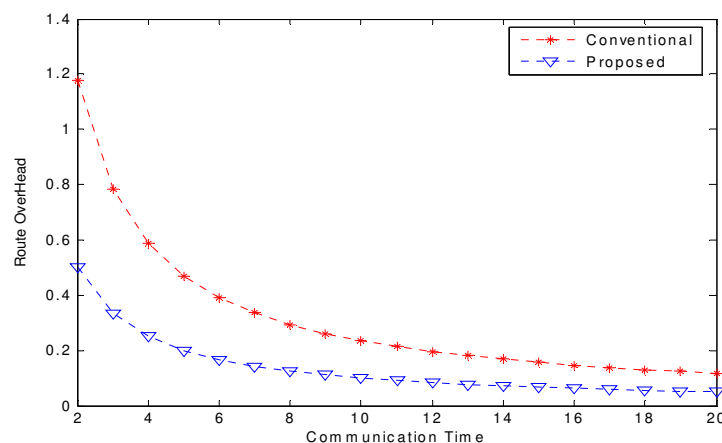


Fig 8: Route overhead

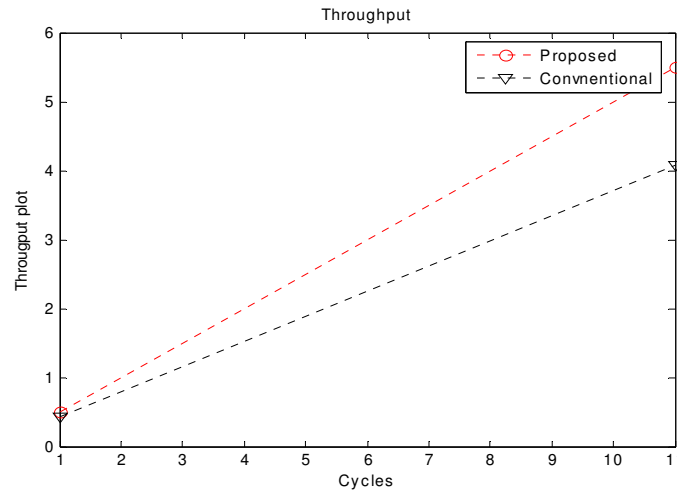


Fig 9: Throughput Plot

V. CONCLUSION

In this paper authors give importance of communication system and based on its increasing usages, need for developments in communication area. In this paper authors identified some problems in optical networks and proposed a new methodology towards developing routing scheme in optical network. So this paper gives a clear idea about a different approach to improve the quality parameters based on adaptive routing mechanism. The concept of route overhead due to queue at the link points is considered. They developed a new model called markovian model, to obtain an optimal routing in optical network so as to achieve the quality service in distributed optical network. The quality metrics developed for the proposed approach is observed to be higher in quality as compared to the conventional approach of routing scheme and finally they explained all these improvements with the simulation results

REFERENCES

- [1]. G. Apostolopoulos, R. Guerin, S. Kamat, and S. Tripathi, "Improving QoS Routing Performance Under Inaccurate Link State Information." Proceedings of the 16th International Tele traffic Congress, June 7-11, 1999.
- [2]. R. A. Barry and S. Subramaniam, "The MAX-SUM Wavelength Assignment Algorithm for WDM Ring Networks," OFC'97, 1997.
- [3]. K. Chan and T.P. Yun, "Analysis of Least Congested Path Routing in WDM Light wave Networks," IEEE INFOCOM'94, vol. 2, pages 962-969, 1994.
- [4]. C. Chen and S. Banerjee, "A New Model for Optimal Routing and Wavelength assignment in Wavelength Division Multiplexed Optical Networks." In Proc. IEEE INFOCOM'96, 1996, pages 164-171.
- [5]. I. Chlamtac, A. Ganz, and G. Karmi, "Purely Optical Networks for Terabit Communication," IEEE INFOCOM'89, pages 887-896, 1989.
- [6]. A. Girard, Routing and dimensioning in circuits switched networks. Addison-Wesley, 1990.
- [7]. H. Harai, M. Murata, and H. Miyahara, "Performance of Alternate Routing Methods in All-Optical Switching Networks," IEEE INFOCOM'97, vol. 2, pages 516-524, 1997.
- [8]. G. Jeong and E. Ayanoglu, "Effects of Wavelength-Interchanging and Wavelength Selective Cross-Connects in Multiwavelength All-Optical Networks," IEEE INFOCOM'96, vol. 1, pages 156-163, March 1996.
- [9]. J. P. Jue and G. Xiao, "An Adaptive Light path Establishment Scheme for Wavelength-Routed Optical Networks," IEEE ICCCN, 2000.
- [10]. L. Li and A. K. Somani, "Dynamic Wavelength routing Using Congestion and Neighborhood Information," IEEE/ACM Transactions on Networking, 1999.
- [11]. E. Karasan and E. Ayanoglu, "Effects of Wavelength Routing and Selection Algorithms on Wavelength Conversion Gain in WDM Optical Networks," IEEE/ACM Transactions on Networking, vol. 6, no. 2, pages 186-196, April 1998.

- [12]. Urmila Bhanja, Sudipta mahapatra, Rajashri Roy, "A novel solution to the dynamic routing and wavelength assignment problem in transparent optical networks", "International Journal of computer networks & communication, Vol.2, No.2, March 2010.
- [13]. Virendra Singh Shekhawat, Dinesh Kumar Tyagi, V.K. Chaubey, "Weight Based Edge Disjoint Path Routing and Wavelength Assignment (WEDP-RWA) Algorithm for WDM Networks", IEEE, 2008.
- [14]. S. Ramamurthy and B. Mukherjee, "Fixed-Alternate Routing and Wavelength Conversion in Wavelength-Routed Optical Networks," IEEE GLOBECOM'98, vol. 4, pages 2295-2302, 1998.
- [15]. S. Subramaniam and R. A. Barry, "Wavelength Assignment in Fixed Routing WDM Networks," IEEE ICC'97, pages 406-410, June 1997.
- [16]. H. Zang, J. P. Jue, B. Mukherjee, "A Review of Routing and Wavelength Assignment Approaches for Wavelength- Routed Optical WDM Networks", Optical Networks Magazine, Vol. 1, No. 1, January 2000, pp 47-60.
- [17]. X. Zhang and C. Qiao, "Wavelength Assignment for Dynamic Traffic in Multi-fiber WDM Networks," IEEE ICCCN, pages 479-485, Oct. 1998.
- [18]. Chan, K. and Yum, T.P., "Analysis of least congested path routing in WDM lightwave networks." INFOCOM '94. Networking for Global Communications, 13th Proceedings IEEE, 1994. pp. 962-969.
- [19]. Banerjee, D. and Mukherjee, B. "A Practical Approach for Routing and Wavelength Assignment in Large Wavelength-Routed Optical Networks." IEEE Journal on Selec. Areas in Comm., vol 14, No. 5, June 1996.
- [20]. Dorigo, M. and Gambardella, L.M., "Ant-Colony System: A Cooperative Learning Approach to the Travelling Salesman Problem." IEEE Transactions on Evolutionary Computation, pp. 53-66.
- [21]. Dijkstra, E. "A note on two problems in connexion with graphs." Numerische Mathematik, 1959. vol. 1, pp. 269-271.
- [22]. Karasan, E. and Ayanoglu E., "Effects of Wavelength Routing and Selection Algorithms on Wavelength Conversion Gain in WDM Optical Networks." IEEE Trans. Networking, vol 6, pp. 186-196, April 1998. Dutton, H.J., Understanding Optical Communications, Prentice Hall, 1999.
- [23]. Hui, Z., Jue, J., and Mukherjee, B. "A Review of Routing and Wavelength Assignment Approaches for Wavelength- Routed Optical WDM Networks," Optical Networks, January 2000.
- [24]. Zhang, X. and Qiao, C. "Wavelength Assignment for Dynamic Traffic in Multi-fiber WDM Networks," ICCCN '98, pp. 479-585, 1998.
- [25]. Stern, T.E. and Bala, K., "Multiwavelength Optical Networks." Addison-Wesley, 1999.

Authors Biography

S. SURYANARAYANA is working as Professor, ECE Department, CMR Institute of Technology, Hyderabad, Andhra Pradesh, INDIA. He received the Bachelors degree in Electronics & communication engineering from JNTU College of Engg in 1991 and M.E(Microwaves) from Birla Institute of Technology (BIT), Ranchi. He is pursuing Ph.D (Optical Communications) under the guidance of Dr. K. Ravindra and Dr. K. Chenna Kesava Reddy. His research interests are Optical Communication, Networking, Switching & Routing and Electromagnetic waves. He has published 12 papers in International/National Journals and Conferences. He is an IEEE Member and life member of ISTE.



K. RAVINDRA currently working as a principal in Mallareddy Institute of Tech & Science, Secunderabad. He received his B.Tech. degree in Electronics and Communication Engineering from Andhra University, M.E. with Honors from University of Roorkee (Presently IIT, Roorkee) and Ph.D. in the field of Mobile Communications from Osmania University. He is a recipient of GOLD MEDAL from Roorkee University for standing first in Masters Degree. He has Engineering teaching experience of over 25 years in various capacities. Dr. Ravindra served as Senior Scientist in the "Research and Training Unit for Navigational Electronics" Osmania University from December 2000 to December 2002. He is presently working as PRINCIPAL in Malla Reddy Institute of Technology & Science, Hyderabad.



Ravindra has 36 technical papers to his credit published in various International, National Journals and Conferences. He co-authored 2 technical reports in the field of GPS. He is the resource person for video lectures produced by SONET, Andhra Pradesh. He is presently guiding 4 research scholars in the fields of Optical & Mobile Communications, Cellular and Next generation networks etc. He is life member of ISTE, IETE and SEMCE(I).

K. CHENNA KESAVA REDDY currently is working as a principal in TKR College of Engg & Tech, Hyderabad. He received his B.E. (Electrical Engineering) in 1973, and M.Tech. (Electronic Instrumentation), from Regional Engineering College, Warangal, 1976. He obtained Ph.D. (Power Electronics), from JNT University, Hyderabad in 2001. He has guided many numbers of graduate and post graduate projects. He has worked in various capacities in JNTU.

ENHANCED BANDWIDTH UTILIZATION IN WLAN FOR MULTIMEDIA DATA

Z. A. Jaffery¹, Moinuddin², Munish Kumar³

¹Associate Professor, Jamia Millia Islamia, New Delhi, India,

²Professor, Jamia Millia Islamia, New Delhi, India,

³Sr. Lecturer, C-DAC, Noida, India,

ABSTRACT

Deployment of wireless local area networks (WLANs) is growing consistently and demanding the support of multimedia applications with acceptable quality of service (QoS). This is attracting the interest of researchers globally. Under the optimum QoS, a number of VoIP calls can be supported by a WLAN. Distributed Coordination Function (DCF) and Point Coordination Function (PCF), two MAC protocols specified in the IEEE 802.11 standard have upper bound on VoIP connections. Under DCF mode 12 calls and in 20 calls in PCF mode [1,2,3,4]. In this paper we are proposing an access media mechanism in which audio data is transmitted in PCF mode and best-effort traffic in DCF mode. In the proposed access media mechanism, polling list is dynamically updated so that only those stations are polled which have voice packets ready to transmit. We have proposed a multi-queued MAC architecture for the access point. We considered voice traffic in CBR mode. The simulation results show that the maximum number of VoIP calls supported by 802.11b is 26 and 14 when inter arrival time for voice packets is 20 ms and 14 ms respectively.

KEYWORDS: Medium Access, Mechanism, Multimedia Data, QoS, Bandwidth.

I. INTRODUCTION

In future generations of WLANs the IEEE 802.11 WLANs will influence the style of daily life of people. The 802.11 technology provides flexible and cheap wireless access capability. Deployment of an 802.11 WLAN is very easy also in hospitals, stock markets, campuses, airports, offices and many other work places. Multimedia applications are increasing very fast in number as well as in size. Demand of voice and broadband video services through WLAN connections is growing day by day. Real time multimedia applications require strict QoS support such as guaranteed bandwidth and bounded delay/jitter etc [6,7,8,17]. In the organization of this paper section-2 describes the functioning of DCF and PCF modes, section -3 discusses the past work done in this direction, section-4 focuses on the limited QoS support in PCF, section-5 describes the proposed approach, in section-6 simulation environment and results are discussed.

II. BACKGROUND

In this section, we are discussing an overview of the IEEE 802.11 standard that provides two different channel access mechanisms, namely the Distributed Coordination Function (DCF) and Point Coordination Function (PCF). Our scheme introduces the enhancements in the PCF access scheme.

2.1 Distributed Coordination Function (DCF)

Most of the wireless LANs in the Industrial Scientific and Medical (ISM) band uses CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) as the channel access mechanism. The basic principles of CSMA are to listen before talk and the contention. This is asynchronous message passing mechanism (connectionless), delivering a best effort of service, and no bandwidth and latency are guaranteed. CSMA is fundamentally different from the channel access mechanisms used by

cellular phone systems (i.e. TDMA).

CSMA/CA is derived from the channel access mechanism CSMA/CD (*Collision Detection*) employed by Ethernet. However, collisions waste valuable transmission capacity, so rather than the collision detection (CD) used in Ethernet, CSMA/CA uses collision avoidance. Collision Avoidance (CA), on a wire, the transceiver has the ability to listen while transmitting and so to detect collisions (with a wire all transmissions have approximately the same strength). But, even if a radio node could listen on the channel while transmitting, the strength of its own transmissions would mask all other signals on the air. Thus, the protocol cannot directly detect collisions like with Ethernet and only tries to avoid them. The 802.11 standard defines the Distributed Coordination Function (DCF) as its fundamental access method and is based on CSMA/CA. DCF allows each multiple independent stations to interact without central control. Figure 1 illustrates the basic access method used in the DCF protocol. If a station finds a channel idle for at least a DIFS period, it sends the first frame from its transmission queue. If channel is busy, the station waits till the end of current transmission and then starts the contention. It selects a random slot time, so called back-off time from a Contention Window (CW) and waits for DIFS and its back-off time.

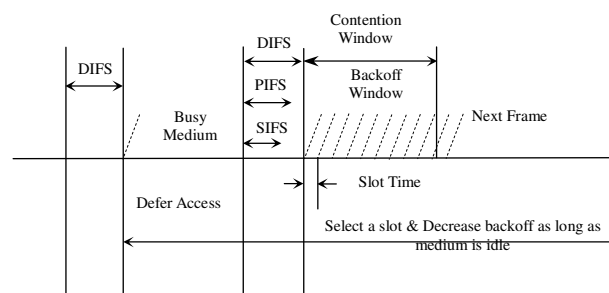


Figure 1: DCF (CSMA/CA) basic access method

The backoff time is calculated as

$$T_{backoff} = \text{Rand}(0, CW) * T_{slot}$$

where T_{slot} is a time slot specific to physical layer and

$\text{Rand}()$ is a uniform distribution random function[2].

The back-off time is computed to initialize the back-off timer and this timer is only decreased when the medium is idle. When the medium is sensed to be busy, this timer is frozen.

When its back off timer expires, and if the channel is still idle, the node sends the frame. Thus, the node having chosen the shortest backoff time wins and transmits its frame. The other nodes just wait for the next contention (after waiting for the end of this packet transmission). Because the contention period is derived from a random number chosen with a uniform distribution, and done for every frame, each station is given an equal chance to access the channel.

2.2 Point Coordination Function (PCF)[10,11,15]

Periods of contention free service arbitrated by the Point Coordinator (PC) alternate with the standard DCF-based access (or contention period). The duration of the contention free period can be configured. 802.11 describes the contention-free period as providing near asynchronous service because the contention-free period will not always start at the expected time.

The Contention-free service uses a centralized access control method. Access to the medium is restricted by the Point Coordinator, a specialized function implemented in access points. Associated stations can transmit data only when they are allowed to do so by the point coordinator. Contention-free access under the PCF resembles token-based networking protocols, with the point coordinator's polling taking the place of a token. Despite, access is under the control of a central entity, all transmissions must be acknowledged. The figure 2 illustrates the PCF access method.

When the PCF is used, time on the medium is divided into contention-free period (CFP) and the contention period (CP). Access to the medium during the CFP is controlled by the PCF, while access

to the medium in CP is controlled by the DCF[12,13,14]. In order to be fair with contending traffic, the contention period must be long enough for the transfer of at least one maximum size frame and its associated acknowledgement. Alternating periods of contention-free service and contention based service repeat at regular intervals, called the contention-free repetition interval (known also as super frame).

At the beginning of the CFP, the PC (which resides in AP) transmits a management frame, called beacon. One of the beacon role components is the maximum duration, *CFPMaxDuration*, of the CFP. The PC generates beacons at regular beacon frame intervals, thus every station knows when the next beacon frame will arrive. This time is called target beacon transmission time (TBTT). All stations receiving the beacon set the NAV to the maximum duration to lock out DCF based access to the wireless medium. The access point maintains a polling list of associated stations and polls any station in this list. Since time in the CFP is precious, acknowledgements, polling, data transfer may be

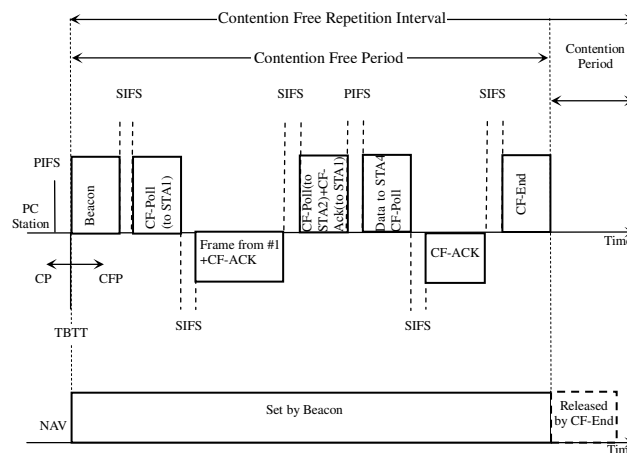


Figure 2: PCF Access Scheme

combined to improve efficiency (as shown in figure 2). All CFP transmissions are separated by short inter frame spaces (SIFS), where PCF waits some time if there is no response from a polled station, to prevent interference from DCF traffic (both are shorter than DCF inter frame space, DIFS).

III. RELATED WORK

In the past many researchers have given analytical, simulation based or experimental results on transmitting Voice over IEEE 802.11 WLAN. MAC protocols defined in IEEE 802.11 standards [1,2,19] are DCF and PCF. Performance of transmitting Voice over WLAN in DCF mode as well as PCF mode has been evaluated. In PCF mode various polling algorithms have been used to transmit Voice over WLAN in the form of the interactive human speech.

In [3] Ali Zahedi and Kevin Pahlavan obtained the capacity of IEEE 802.11 WLAN with voice and data services analytically in DCF mode. Precisely the question which they answered is the number of network telephone calls that can be carried over WLAN with a predefined amount of data traffic or what is the maximum data traffic per user for a given number of voice users. Priority to voice is given by assigning UDP protocol for voice and TCP protocol for data. They found that in 1Mbps bandwidth, a maximum 18 voice users are supported with upper bound for delay of 100ms and decreasing upper bound for delay to 50ms which reduces maximum voice users to 14. Data traffic is assumed to be less than 10Kbps.

Sachin Garg and Martin Kappes found an upper bound on the number of simultaneous VoIP calls that can be placed in a single cell of an 802.11b network in [4]. They performed an experiment in which multiple Wireless PCs running Windows 2000, were associated with the same 802.11b AP, which was connected to a 100Mbps Ethernet. The setup was used to make full duplex VoIP calls between a wireless PC and a wired PC using IP phones. For each call the ITU G.711 codec was used where frames are sent every 10ms. Each call results in two RTP streams, from wired to wireless and vice-versa. Number of VoIP connections with acceptable voice quality is tested by successively establishing new calls in addition to the ongoing calls. The quality of connections was monitored

measuring of loss, jitter and round trip time by a commercially available tool. For the first five calls, the quality of all the calls was acceptable. Loss (0%), round trip time (around 5ms) and jitter (around 7ms) were all in acceptable ranges for a good quality of VoIP. When the sixth call was placed, except for an increase in the round-trip time for some of the connections the quality of all six simultaneous connections was still acceptable. As soon as the seventh call was placed, all seven wired to wireless streams started suffering approximately 16% loss and the call quality became unacceptable for all calls in this direction. All wireless to wired streams still exhibited acceptable quality. In addition to this experiment they obtained an upper bound for simultaneous calls analytically. They found that when a G.711 Codec with 20ms audio payload is used, an 802.11b cell could support only 3 to 12 simultaneous VoIP calls. The actual number depends on the effective transmission rate of the wireless station, which for 802.11b can be 1Mbps, 2Mbps, 5.5Mbps and 11Mbps.

In [6,16] practical investigation of the IEEE 802.11b MAC layer's ability to support simultaneous voice and data applications is done. DCF mechanism is modified and new mechanism is called Back-off Control with Prioritized Queuing (BC-PQ). BC-PQ addresses the two shortcomings of the DCF mode with respect to voice. First, it distinguishes voice packets from data packets and provides a higher priority to the voice traffic. Allocating separate prioritized queues for voice and non-traffic does this. Secondly, in addition to priority queuing, the enhanced AP transmits voice packets using zero back-off instead of random back-off as required by the 802.11b standard. The key parameter used to quantify voice performance is packet loss.

Jing-Yuan Yeh and Chienhua Chen in [7] proposed three polling schemes (RR, (FIFO, Priority and Priority Effort-Limited Fair) combined with the Point Coordination Function (PCF) to improve the utilization of the wireless channel and support certain Quality of Service of multimedia traffic. The polling schemes proposed in [8] are-Round-Robin Scheme (RR), First-In-First-Out Scheme (FIFO), Priority Scheme, Priority Effort-Limited Fair Scheme. All the above-mentioned polling schemes are simulated with the network simulator OPNET in [7]. It is found through simulations that all these schemes perform better than the DCF mode. To achieve the maximum throughput in a BSS, the FIFO scheme is found to be the best. Priority Scheme provides a simple way to support QoS of traffic; however this scheme can exhaust all the bandwidth of best-effort traffic. The Priority-ELF scheme achieves high utilization of wireless channel in case of the bursty traffic.

In [9,10,15] the capability of the Point Coordination Function (PCF) to support Voice over IP (VoIP) applications is evaluated. The capability of PCF mode in support of variable bit rate (VBR) VoIP traffic is investigated, where the silence suppression technique is deployed in voice codec so that no voice packets are generated in silence periods. Simulation shows that under the PCF using VBR mode for the VoIP traffic may effectively reduce the end-to-end delay of VoIP. Simulation is carried out in the OPNET network simulator. The upper bound for number of VoIP connections in CBR mode is found to be 15 and in VBR mode it is 22. Brady's model and May and Zebo's models are used for VBR voice traffic.

E. Ziouva and T. Antonakopoulos in [12] proposed a new dynamically adaptable polling scheme for efficient support of voice communications over IEEE 802.11 networks. They proved analytically that when silence detection is used their scheme improves the capability of IEEE 802.11 wireless LANs for handling voice traffic efficiently. Their polling scheme is called Cyclic Shift and Station Removal Polling Process (CSSR).

In [11] a distributed fair queuing scheme called Distributed Deficit Round Robin (DDRR) is proposed which can manage bandwidth allocation for delay sensitive traffic. The question which is answered in this paper is how many voice and video connections can be accommodated in an 802.11 WLAN satisfying the imposed QoS requirements under different scheduling schemes.

Xiyan Ma, Cheng Du and Zhisheng Niu in [14] proposed two adaptive polling list arrangement schemes called Dynamic Descending Array (DDA) and Hybrid DDA and RR (HDR) in order to decrease the average delay of the voice packets, by means of reducing the possibility of null polls. One scheme is called Dynamic Descending Array (DDA) and the other scheme is a combination of DDA and traditional Round-Robin scheme and is called Hybrid DDA and RR (HDR).

IV. PCF WITH LIMITED QoS SUPPORT

Although the contention-free service is designed in 802.11 networks to provide QoS for real-time

traffic, this service also have some limitations[8,9]. In the following we describe main limitations related PCF-

- **Unpredictable beacon delay**-The problem is related to the uncontrolled length of CP. Indeed, the minimum length of the CP is the time required to transmit and acknowledge one maximum frame size. It is possible for the contention service to overrun the end of the CP, due to transmission of a contending traffic. When the contention based service runs past the TBTT, the CFP is foreshortened, and hence the beacon is delayed.
- **Unknown transmission time of polled stations**- A station which is polled by the PC is allowed to send a single frame that may be fragmented and of arbitrary length, up to maximum of 2304 bytes (2312 bytes with encryption). Furthermore, different modulation and coding schemes are specified in 802.11a, thus the duration of the MSDU delivery that happens after the polling is not under the control of PC. This may destroy any attempt to provide any QoS to other stations that are polled during the rest of the CFP.
- **No Knowledge of the offered traffic at the stations**- With CFP, the access point (AP) has any knowledge of the offered traffic at the polled stations. Thus, when polling the different stations with a round-robin scheduling algorithm, the PC may waste a lot of time until polling a special station having critical time traffic (e.g., CBR traffic). This may affect the QoS parameters for these traffic categories. Hence, with PCF there is no efficient scheduling algorithm, which has the knowledge of the different traffic categories at associated stations and uses this knowledge in order to meet the requirements (e.g., latency, bandwidth) of these different traffic categories.

V. PROPOSED MECHANISM TO ACCESS MEDIA

5.1 Overview

In this section we will discuss our proposed access media mechanism to transmit voice over WLAN. In our scheme also we use PCF mode to transmit voice and DCF mode to transmit Best-Effort traffic. A lot of valuable bandwidth is wasted when a station which has nothing to transmit is polled. Keeping all the stations in polling list and polling them in Round-Robin manner hampers the channel utilization severely. We propose a dynamic polling list arrangement to poll only those stations which have data to transmit. Also in our scheme, the downlink voice traffic from AP is combined with CF-Poll which can significantly reduce the overhead and thus may improve the channel utilization. We give the MAC architecture of a WLAN station and the MAC architecture of the access point respectively. How the polling list is handled dynamically is discussed in the section 5.4.

5.2 MAC Architecture of WLAN Station

MAC architecture of a WLAN station in our scheme is shown in the figure 3. Each station maintains two queues at its MAC layer, one for voice traffic and the other for Best-Effort traffic. When a frame arrives at the MAC layer of a WLAN station from its upper layers a classifier checks whether it is a voice frame or Best-Effort data frame, and accordingly it is put in the voice queue or Best-Effort queue. When a station is polled in the CFP it transmits the voice frame from the head of the voice queue. The frames from the Best-Effort queue are transmitted in DCF mode in the CP.

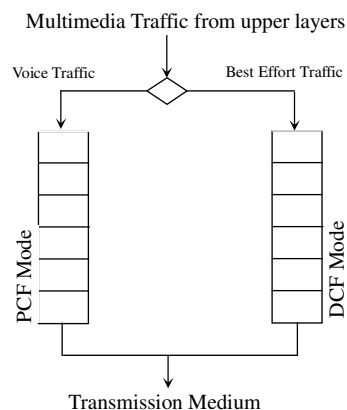


Figure 3: MAC Architecture of Station

5.3 MAC Architecture of the Access Point

MAC architecture of the access point, in our scheme is shown in the figure 4. The AP maintains one queue for Best-Effort traffic and (Poll Limit + 1) queues for voice traffic at its MAC layer. Poll Limit is the maximum number of stations which can be kept in the polling list in our Dynamic Polling arrangement scheme discussed in 4.4. When a frame arrives at the MAC layer of the AP, a classifier checks whether it is a voice frame or data frame and accordingly it takes its path as shown in the figure 4. If it is a data frame, it simply enters the Best-Effort queue and is transmitted in CP using DCF method. Otherwise, if it is a voice frame, further it is checked whether the destination station is in polling list or not. In case destination station is not in the polling list, it is checked whether number of stations in polling list exceed Poll_Limit or not. If they exceed then the frame enters the Non_Poll queue, otherwise destination station is added in the polling list and the frame takes the path as shown in the figure 4. The frame enters the Non_Poll queue as shown in figure 4. On the other hand, if the destination station is in the polling list, then the frame enters queue for its destination station among Poll_Limit queues.

5.4 Dynamic Polling list arrangement

The main issues in maintaining a polling list are how to add a station in the polling list and how to remove a station from the polling list.

5.4.1 Adding a station in the polling list:

The stations which want to transmit voice frames in CFP send Resource Reservation (RR) request in CP in controlled contention interval (CCI). A CCI is started when the PC sends a specific control frame. Only those stations which want to transmit voice in the CFP contend in CCI using CSMA/CA. In case of collision between RR requests no retransmission is done. Corresponding stations can send their RRs in the next CCI. If a downlink frame arrives at the MAC layer of the AP and the destination station is not in the polling list, in this scenario if number of stations in the polling list is less than Poll Limit then the destination station is added to the polling list.

5.4.2 Removing a station from the polling list:

A station is removed from the polling list when its connection time is over.

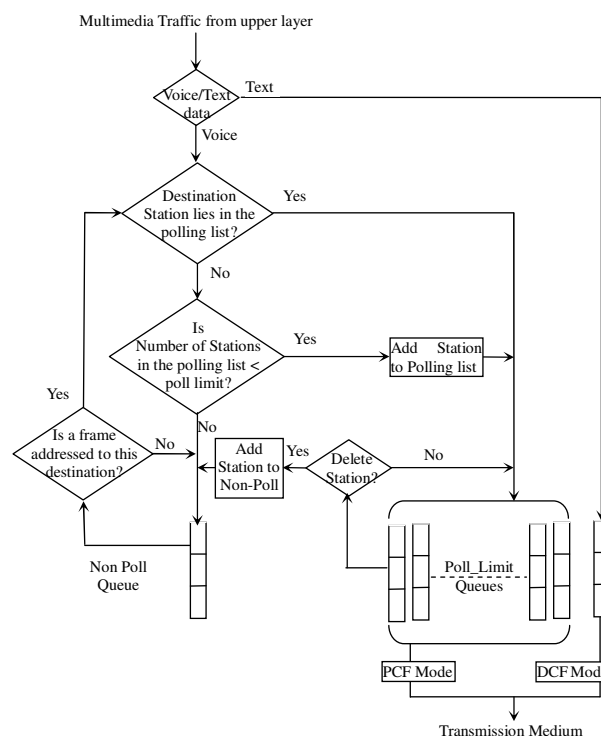


Figure 4: MAC Architecture of a Polling List

VI. SIMULATION STUDIES ON THE PROPOSED ACCESS MEDIA MECHANISM

6.1 Overview

We have simulated the proposed access media mechanism in C programming language. All the parameters considered for simulation are also given in the table 1. Finally simulation results are discussed in the section 5.3. Parameters used for simulation are shown in the table 1. The simulation is run for 85000 cycles of contention-free repetition interval. The results of initial warm-up period of 5000 cycles are ignored.

6.2. Maximum Number of VoIP Connections in 802.11b

Calculation shows that given the 11 Mbps rate and 128kpbs needed for duplex VoIP calls. But due to network layer protocol, inter frame space, beacon frame, poll frame, CF-end frame and channel utilization is around 30% of the total payload. Figure 5 shows the plot between number of VoIP connections and average packet delay when the inter arrival time is 20ms. We see in that case the maximum number of voice calls is 26; the uplink average packet delay goes up abruptly. We see when there is no best-effort traffic the maximum number calls supported by 802.11b is 26.

Table 1: Simulation Parameter

PHY Layer Specifications	DSSS
Transmission Rate	11 Mbps
Beacon Interval	60ms
Beacon Size	106 bytes
QoS Acknowledgement	14 bytes
CF-Poll Frame Size	34 bytes
CF-End Frame Size	20 bytes
PLCP Preamble	18 bytes
PLCP Header	6 bytes
SIFS Time	10 μ s
PIFS Time	30 μ s
DIFS Time	50 μ s
A Slot Time	20 μ s
Nccop	5
RTP Header	8 bytes
UDP Header	20 bytes
IP Header	8 bytes
MAC Header	34 bytes
CWmin	7
CWmax	255
Retry limit	5

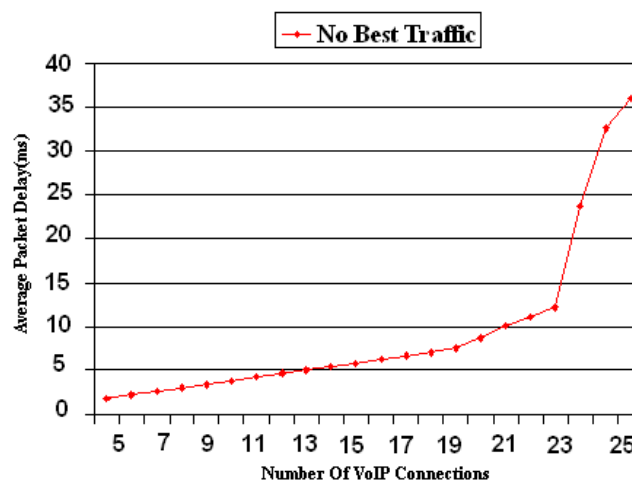


Figure 5. Number of VoIP vs. Average Packet Delay
(Inter Arrival Time is 20 ms)

6.2 Simulation Results

In another case, Figure 6 shows the plot between number of VoIP connections and average packet delay when the inter arrival time is 10ms. We see in that case the maximum number of voice calls is 14; the uplink average packet delay goes up abruptly. We see when there is no best-effort traffic the maximum number calls supported by 802.11b is 14.

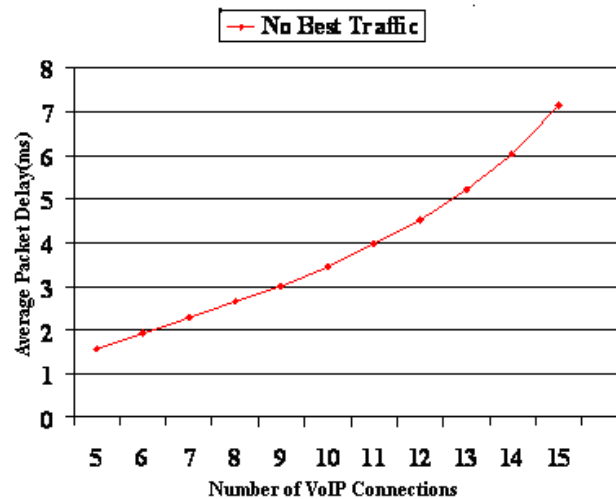


Figure 6. Number of VoIP vs. Average Packet Delay
(Inter Arrival Time is 10 ms)

As the best-effort traffic increases up to 10% the numbers of voice calls in this case are 24 and 13. This is shown in figure 7 and figure 8.

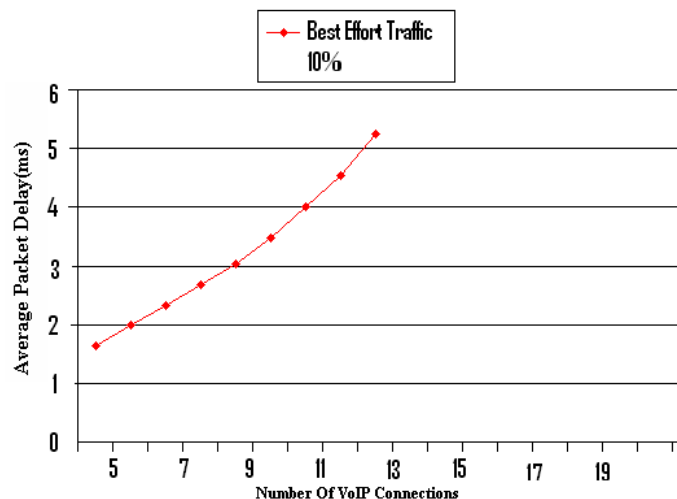


Figure 7. Number of VoIP Connections vs. Average Packet Delay
(Inter arrival Time is 20 ms)

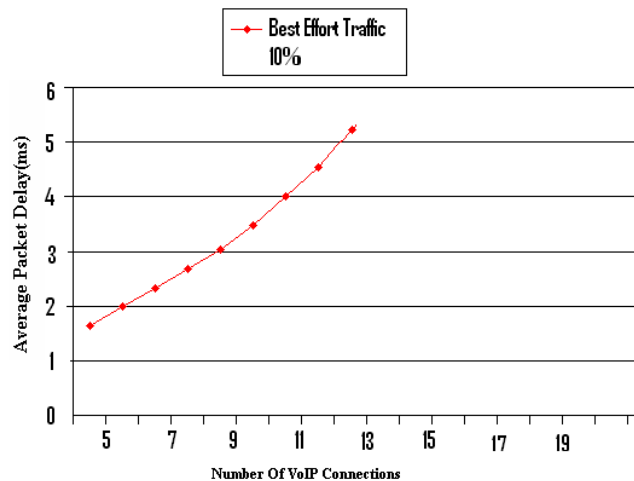


Figure 8. Number of VoIP Connections vs. Average Packet Delay
(Inter arrival Time is 10 ms)

VII. CONCLUSION

In this work we studied the VoIP over the 802.11 networks, for the perspective of number of the connections that an access point can support. We found that the maximum number of full duplex VoIP connections supported by 802.11b is 26. Our access media mechanism improves on the maximum number of VoIP connections supported. This enhancement is due to efficient utilization of the available bandwidth hence will support to multimedia real time applications as well.

REFERENCES

- [1]. IEEE Std 802.11, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. 1999.
- [2]. IEEE Draft Std 802.11e, *Amendment: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*, D2.0a. Nov. 2001.
- [3]. Ali Zahedi and Kevin Pahlavan, "Capacity of a Wireless LAN with Voice and Data Services," *IEEE Transactions on Communications*, Vol.48, no.7, pp.1160-1170, July 2000.
- [4]. S. Garg and M. Kappes, "Can I add a VoIP call?" *IEEE International Conference on Communication*, pp.779-783, 2003.
- [5]. F. Anjum, M. Elaud, D. Famolari, A. Ghosh, R. Vidyanathan, A. Dutta and P. Aggarwal, "Voice Performance in WLAN Networks-An Experimental Study," *IEEE Globecom*, 2003.
- [6]. M. Veeraraghavan, N. Cocker and T. Moors, "Support of Voice Services in IEEE 802.11 Wireless LANs," *IEEE INFOCOM'01*, vol. 1, pp. 448-497, April 2001.
- [7]. Jing-Yuan Yeh and Chienhua Chen, "Support of Multimedia of Services with the IEEE 802.11 MAC protocol," *IEEE International Conference on Communication*, 2002.
- [8]. Andreas Kopsel and Adam Wolisz, "Voice Transmission in an IEEE 802.11 WLAN Based Access Network," *Wow Mom, Rome Italy*, pp. 24-33, July 2001.
- [9]. D. Chen, S. Garg, M. Kappes and Kishor S. Trivedi, "Supporting VBR VoIP Traffic in IEEE 802.11 WLAN in PCF mode," *Opnet Work 2002, Washington DC*, August 2002.
- [10]. E. Ziouvra and T. Antonakopoulos, "Efficient Voice Communications over IEEE 802.11 WLANs Using Improved PCF Procedures," *The Third International Network Conference-INC 2002*, July 2002.
- [11]. Xian Ma, Cheng Du and Zhisheng Niu, "Adaptive Polling List Arrangement Scheme for Voice Transmission with PCF in Wireless LANs," *10th Asia-Pacific Conference on Communications and 5th International Symposium on Multi-Dimensional Mobile Communications*, 2004.
- [12]. Ravindra S. Ranasinghe, Lachlan L.H. Andrew and David Everett, "Impact of Polling Strategy on

- Capacity of 802.11 Based Wireless Multimedia LANs,” *IEEE International Conference on Networks, Brisbane Australia*, 1999.
- [13]. T. Kawata, S. Shin, Andrea G. Forte and H. Schulzrinne, “Using Dynamic PCF to Improve the Capacity for VoIP Traffic in IEEE 802.11 Networks,” *IEEE Wireless Communications and Network Conference*, March 2005.
- [14]. D.Chen, S. Garg, M. Kappes and Kishor S. Trivedi, “Supporting VoIP Traffic in IEEE 802.11 WLAN with Enhanced MAC for Quality of Service,” *Opnet Work 2003, Washington DC*, September 2003.
- [15]. Md. Atiur Rahman Siddique and Joarder Kamruzzaman, “Performance Analysis of PCF based WLANs with Imperfect Channel and Failure Retries”, *GLOBECOM 2010, 2010 IEEE Global Telecommunications Conference, Miami, FL*, Dec-2010, pp 1-6.
- [16]. Suchi Upadhyay, S.K.Singh, Manoj Gupta, Ashok Kumar Nagawat, “Improvement in Performance of the VoIP over WLAN”, *International Journal of Computer Applications (0975 – 8887) Volume 12– No.4*, December 2010, pp 12-15.
- [17]. Bum Gon Choi, Sueng Jae Bae, Tae-Jin Lee and Min Young Chung, “Performance Analysis of Binary Negative Exponential Backoff Algorithm in IEEE 802.11a under erroneous Channel Conditions”, *ICCS Part-II, LNCS 5593*, pp 237-249, 2009.
- [18]. Suparerk Manitpornsut, Bjorn Landfeldt, “On the Performance of IEEE 802.11 QoS Mechanisms under Spectrum Competition”, *IWCMC’06*, July 2006, *Vancouver, British Columbia, Canada*, , pp 719-724.
- [19]. Dileep Kumar, Yeonseung Ryu, Hyuksoo Jang “Quality of Service (QoS) of Voice over MAC Protocol 802.11 using NS-2”, *CommunicabilityMS’08*, October, 2008, *Vancouver, BC, Canada*, pp 39-44.

Authors

Z. A. Jaffery-He obtained his B. Tech and M. Tech in Electronics Engineering from Aligarh Muslim University, Aligarh, India in 1987 and 1989 respectively. He obtained his PhD degree from Jamia Millia Islamia (a central Govt. of India university) in 2004. Presently he is the associate professor in the Department of Electrical Engineering, Jamia Millia Islamia, New Delhi. His research area includes Application of soft computing Techniques in Signal Processing, Communication engineering and Computer Networking.



Moin Uddin-He obtained his B. Tech and M. Tech in Electrical Engineering from Aligarh Muslim University, Aligarh, India in 1972 and 1978 respectively. He obtained his PhD degree from university of Roorkee in 1992. Dr. Moinuddin is the professor in the Department of Electrical Engineering, Jamia Millia Islamia, New Delhi. Presently he is on deputation as Pro-Vice Chancellor of Delhi Technological University, New Delhi. He has guided several PhD. His research area includes computer networking, soft computing and Artificial Intelligence



Munish Kumar-He obtained his B.E. in Computer Science & Engg from MNREC, University of Allahabad, in 1992 and Master of Computer Science & Engg from Jadavpur University, Kolkata. Tech. in 1995. He is working as Astd. Professor in School of IT, CDAC, Noida. Presently is pursuing his PhD from Jamia Millia Islamia. His research area includes Ad-hoc Networks, Sensor Networks, Wireless Networks, Mobile Computing, Real-Time Applications.



ANALYSIS AND INTERPRETATION OF LAND RESOURCES USING REMOTE SENSING AND GIS: A CASE STUDY

S.S. Asadi ¹, B.V.T.Vasanth Rao², M.V. Raju ³ and M.Anji Reddy⁴

¹Assoc. Prof., Deptt. of Civil Engineering., KL University, Green fields, Vaddeswaram,
Guntur, A.P, India

²Asstt. Prof., Deptt. of Civil Engineering., P.V.P. Siddhardha Institute of Tech.,
Kannure, Vijayawada.

³Asstt. Prof., Deptt. of Civil Engineering, Vignan University, Vadllamudi, Guntur, A.P, India

⁴Director Foreign Relations, Jawaharlal Nehru Technological University,
Hyderabad, A.P, India

ABSTRACT

The human Activities are Constantly adding industrial, domestic and Agricultural wastes to the ground water reservoirs at an alarming rate . In the last few decades, parallel with rapidly developing technology, increase in population and urbanization we have been witnessing alarmed phenomena all over the world. Anthropogenic activities including generation and indiscriminate disposal of solid wastes and extensive use of fertilizers have resulted in increasing levels of air, water and soil pollution, changing land use patterns, decrease in arable land and other dominant problems. The thematic map of the study area is prepared from linearly enhanced fused data of IRS-ID PAN and LISS-III merged satellite imagery and Survey Of India (SOI) toposheets on 1:50,000 scale using visual interpretation technique and using AutoCad and Arc/Info GIS software forming the spatial database.

KEYWORDS: Thematic maps, groundwater quality, remote sensing and GIS

I. INTRODUCTION

Man needs Land for domestic purposes such as cooking, cleaning utensils, gardening, washing clothes and above all for drinking. It is also needed for commercial, industrial and recreational purposes. Land used for such purposes should be not polluted, but should be of good quality. Urbanization and industrialization have directly or indirectly polluted most of the land sources on a global scale. Impact studies can contribute to improve urban development and environmental planning at the project and policy levels and it also introduces analytical tools to support such planning. Remote sensing applications have been operationalized in most of the natural resource management themes and at present the trend is on integrated surveys to arrive at sustainable developmental packages. Keeping this in view, an attempt is made.

II. STUDY AREA

The Maripeda Mandal lies geographically between latitudes 17° 20' 00'' and 17° 35' 00'' and longitudes 79°45' 00'' to 80°00' 00'' is covered in the Survey of India toposheet numbers 56 O/14 and 56 O/15. It is one of the 51 Mandals of Warangal district, in Andhra Pradesh. Maripeda town is at a distance of 90 kms. from Warangal (District H.Q.) and 120kms from Hyderabad (State Capital).

2.1 Study Objectives

- To prepare the digital thematic maps namely Base map, Transport network map, Geomorphology map, Ground Water Potential map, Land use/ Land cover, Hydro geomorphology, Physiographic map, Waste land map, Drainage map etc. using satellite data, collateral data and field data on ARC/INFO GIS platform. This constitutes the spatial database.

III. METHODOLOGY

3.1 Data collection

Different data products required for the study include Survey of India (SOI) toposheets bearing with numbers 56O/14 and 56O/15 on 1:50,000 scale. Fused data of IRS-ID PAN and LISS-III satellite imagery obtained from National Remote Sensing Agency (NRSA), Hyderabad, India. Collateral data collected from related organizations, comprises of water quality and demographic data [2].

3.2 Database creation

Satellite imageries are geo-referenced using the ground control points with SOI toposheets as a reference and further merged to obtain a fused, high resolution (5.8m of PAN) and colored (R,G,B bands of LISS-III) output in EASI/PACE Image processing software. The study area is then delineated and subsetting from the fused data based on the latitude and longitude values and a final hard copy output is prepared for the generation of thematic maps using visual interpretation technique as shown in Figure 1. These thematic maps (raster data) are converted to vector format by scanning using an A0 flatbed deskjet scanner and digitized using AutoCAD software for generation of digital thematic maps using Arc/Info and ARCVIEW GIS software. The GIS digital database consists of thematic maps like land use/land cover, drainage, road network using Survey of India (SOI) toposheets and fused data of IRS - ID PAN and IRS-ID LISS-III satellite imagery (Figure 2).

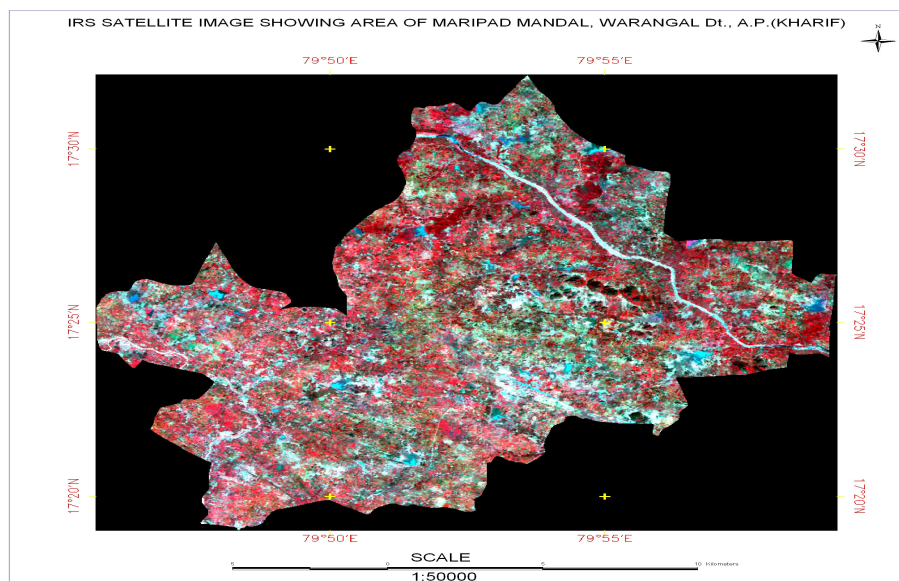


Figure1. Fused satellite imagery

3.2.1 Spatial Database

Thematic maps like base map and drainage network maps are prepared from the SOI toposheets on 1:50,000 scale using AutoCAD and Arc/Info GIS software to obtain a baseline data. Thematic maps of the study area was prepared using visual interpretation technique from the fused satellite imagery (IRS-ID PAN + IRS-ID LISS-III) and SOI toposheets along with ground truth analysis. All the maps are scanned and digitized to generate a digital output (Figure 1).

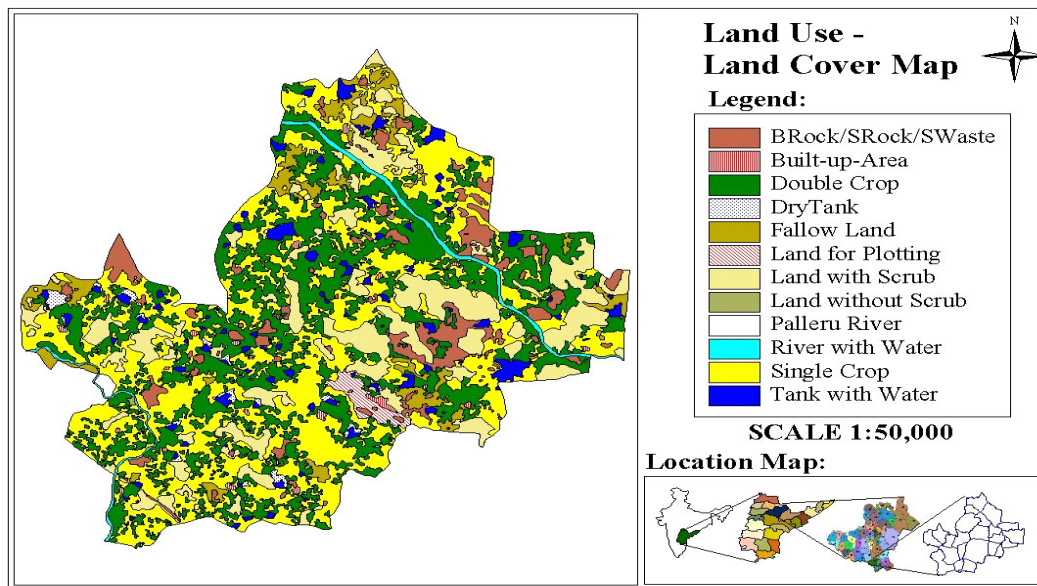


Figure 2 .Land Use/Land Cover

IV. RESULTS AND DISCUSSION

4.1 Village Map

This shows the geographical locations of all villages in the Mandal is called Village Map [1]. This map is prepared by digitization of the maps of Central Survey Office. Revenue boundaries of all the villages are plotted in this map. The entire data of the village available is attached to this Map as database using GIS. This database is very useful to know the present scenario of a village. This map is used for analyzing village wise land resources. In the study area there are 23 revenue villages, out of these Maripeda, which is the Mandal head quarter. By preparing the Village map feature of the individual village can be easily identified.

4.2 Base map

It consists of various features like the road network, settlements, water bodies, canals, railway track, vegetation etc. delineated from the toposheet. The map thus drawn is scanned and digitized to get a digital output. The information content of this map is used as a baseline data to finalize the physical features of other thematic maps.

4.3 Transport Network Map

In the study area all the settlements are connected either by Metalled road or Un-Metalled road. Where as, State Highway connects Maripeda. Railway network does not exist in the Maripeda Mandal. The nearest railway station is Khammam, which is at a distance of 18kms SouthEast of Maripeda village.

4.4 Drainage

Drainage map is prepared by using Survey of India Topographic maps on 1:50,000. All the streams and tanks existing in the study area are marked in this map. These streams further classified based on stream ordering. Only two minor rivers namely Palleru and Akeru exists [3]. The drainage system existing is dendritic. Tank bunds are also marked in the map.

4.5 Watershed characteristics:

The watershed map is prepared in accordance with the National Watershed Atlas and River Basin Atlas of India, 1985. According to this, India is divided into 6 regions out (watershed Atlas of India, 1990) of which the present study area comes under Region-4 and part of basins D, catchment 1, sub-

catchment C,D. The study area is under sub-catchment C watersheds (4D1C6,4D1C7) and under D sub-catchment. (4D1D2) are coming under these watersheds 71 sub-watersheds has delineated [4] .

4.6 Slope map

Slope classes 1, 2 and 3 are observed in the study area. Most of the study area is covered by nearly level, very gently, gently slope class (92%). Small part of the study area (4%) comes under moderately sloping class 4 and (2%) study area comes under the strongly sloping class 5 (IMSD Technical Guidelines, 1995).

4.7 Land Use/Land Cover

The land use/land cover categories such as built-up land, agriculture, forest, water body and wastelands have been identified and mapped from the study area (Figure 3). Major part of the study area is covered with single crop and double crop (93%). About (0.015%)of the study area is under built-up land and Industrial area is(0.017%). From the satellite data the agriculture area (96.05%) could be clearly delineated as four categories, single crop, double crop, fallow land and plantations [5]. Though single crop and double crop has been observed at various parts of the study area and plantations are observed at some places of the study area. Water bodies occupied(0.18%). About (0.46%) of the study area is under scrub forest and(4.21%) of area is under wasteland. Under this category land with scrub (3%), land without scrub (0.24%) and barren sheet rock area (0.09%) are observed (Figure 3).

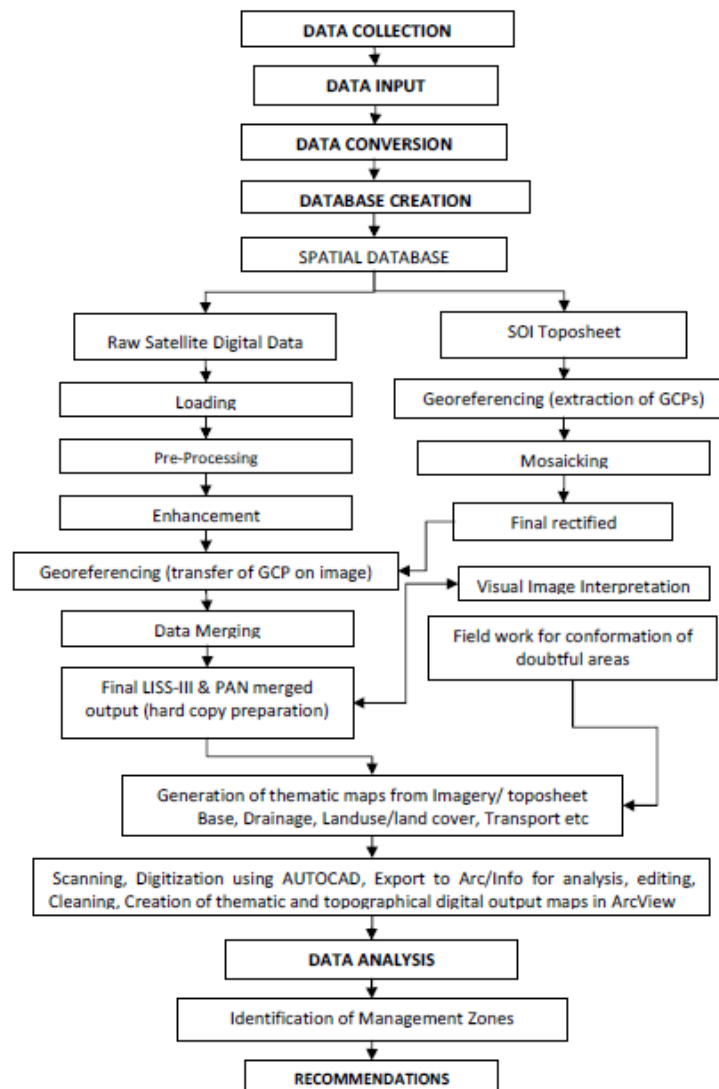


Figure 3: Flow chart showing the methodology adopted for the present study

4.8 Geomorphology

The geomorphological classes observed in the study area are Pediplain with moderate weathering (PPM) (42%), Pediplain with shallow weathering (PPS) (31%), valley (v) (14%), pediment (PD) (8%), pediment inselberg complex (PIC) (2%), inselberg (1%), pediment (1%) and dyke and dyke ridge (0.12%) .

4.9 Geology:

The study area constitutes mainly a granitic terrain (pink-grey) exposing a variety of archaean granitoides of peninsular gneissic complex (PGC) and schistose (older metamorphic) rocks. They are intruded by basic dykes (Proterozoic) and covered locally by the Deccan traps (upper Cretaceous to lower Eocene) [6] . The geological categories observed in the study area are mainly granite (98%), basalt (2%), and some of lineaments, dolerites and pegmatites .

4.10 Soil

The specific objectives of soil mapping are identification, characterization and classification of the soils of the area. The soil types identified in the study area are (1) loamy-skeletal, mixed, rhodic paleustalfs (55%). (2) Fine loamy, mixed, fluventic ustropepts (10%) (3) Fine, montmorillonitic, typic haplusterts (35%).

4.11 Groundwater potential

The groundwater potential map is prepared based on the analysis of various themes such as geomorphology, land use / land cover, lineament, intersection points, drainage pattern, lithological evidences by using converging evidence concept, besides the collateral data obtained from State Groundwater Board with necessary field checks [7] . The groundwater potential map reveals the available quantum of groundwater and is delineated into zones showing high (53%), medium (30%), low (17%), groundwater potential areas.

V. CONCLUSIONS AND RECOMMENDATIONS

1. Through the analysis of soils attribute data, it is clear that 39% of the study area is affected by erosion. In future this may lead to sedimentation and other consequential problems to the major water bodies of study area. This could be best controlled by construction of gully control bunds and extensive reforestation or through agricultural soil conservation and management practices.

2. As irrigation water requirement varies with different crops, cropping pattern in the study area is to be changed for optimum utilization of this resource. Crops like pulses, vegetables should be cultivated, which may result in the reduction of water requirement as well as fertilizer and pesticide load in the study area.

3. The three key activities that are essential for the development of a watershed area are,

- Irrigation management
- Catchment management
- Drainage basin monitoring and management

To address these three activities planners need physical characteristic information on comprehensive lines. Hence, the present work, concentrated on the development of physical characteristics for this study area. The planners at execution level can rely upon such kind of physical characteristic information system for various other watersheds.

4. This study has been concluded to above stated findings. But this study will be useful as input base line data for models like LWAT (Land and Water Assessment Tool) that give more precise and detailed long term predictions on land and water resources.

REFERENCES

- [1] District Census Handbook of Hyderabad, 1991 Directorate of Census Operations, Andhra Pradesh, Census of India.
- [2] APHA, AWWA, WPCF, 1998 Standard Methods for the Examination of Water and Wastewater. (20th edition). American Public Health Association, Washington DC, New York.
- [3] Tiwari, T.N and Mishra, M, 1985 A preliminary assessment of water quality index to major Indian rivers. Indian Journal of Environmental Protection, 5(4), 276-279.
- [4] Mahuya Das Gupta Adak, Purohit KM, Jayita Datta, 2001 Assessment of drinking water quality of river Brahmani. Indian Journal of Environmental Protection, 8(3), 285-291.
- [5] Pradhan, S.K., Dipika Patnaik and Rout, S.P, 2001 Water quality index for the ground water in and around a phosphatic fertilizer plant. Indian Journal of Environmental Protection, Vol.21, 355-358.
- [6] Srivastava, A.K., and Sinha, D.K, 1994 Water Quality Index for river Sai at Rae Bareli for the premonsoon period and after the onset of monsoon. Indian Journal of Environmental Protection, Vol.14, 340-345.
- [7] Kurian Joseph, 2001 An integrated approach for management of Total Dissolved Solids in reactive dyeing effluents. Proceedings of International Conference on Industrial Pollution and Control Technologies, Hyderabad.

Authors

A. Sivasankar is working as Assoc. Prof., Deptt. of Civil Engineering., KL University, Green fields, Vaddeswaram, Guntur, A.P, India. He has 14years of research experience & supervised 2 M.Sc & 2 M.Tech Dissertation. He was Principal Investigator **DST Sponsored Fast Track Young Scientist** Project cost of Rs**15,24,000/-**.



IPV6 DEPLOYMENT STATUS, THE SITUATION IN AFRICA AND WAY OUT

Agbaraji E.C., Opara F.K., and Aririguzo M.I.

Electrical Electronic Engineering Deptt., Federal University of Technology Owerri, Nigeria

ABSTRACT

The number of internet connected devices is increasing terrifically, with each device assigned a unique Internet Protocol (IP) address at a time. Hence the expected problem of IPv4 address exhaustion in the near future called for a better and permanent solution, which is switching to IPv6. Adoption and deployment of IPv6 recorded a fast growth rate globally, though the slow growth rate recorded in Africa was suspected to be due to the poor capacity building and the level of the IPv6 awareness campaign in the region. It was concluded that the developmental strategies created to help in the deployment of IPv6, such as the global awareness campaign, was confirmed effective. Also the World IPv6 day provides a 24 hours experiment to uncover the challenges of the transition to IPv6 and to develop measures to resolve them.

KEYWORDS: Internet, IPv4, IPv6, RIRs, AfriNIC, ISP.

I. INTRODUCTION

Today, most electronic devices such as mobile phones, Personal Digital Assistants (PDAs), PCs, Internet telephones, etc use in homes and other places, rely on the internet technology for their various services. The internet connected devices use the internet protocol (IP) address to communicate over the network with each device assigned a unique IP address. This means that, for any device to communicate through the internet, it must be assigned an IP address. Most private and business application services (online transactions), including social activities such as Facebook, Twitter, Yahoo, etc., depend on the IP address for their functions. Thence, the tremendous growth rate in the number of internet connected devices and high dependence on the internet for human daily activities have caused the expected exhaustion of the long-time used IPv4 addresses.

The two versions of IP currently in use; Internet Protocol Version Four (IPv4) and Internet Protocol Version Six (IPv6), with the IPv6 adopted proactively to solve the expected problem of the first and widely used version (IPv4) exhaustion in the future. The Number Recourse Organization (NRO), made up of five regional internet registries (RIRs) was set up to work together at global and regional levels to promote the transition from IPv4 to IPv6 and layout strategies to manage the distribution of the remaining unallocated IPv4 address pool [2].

The objective of this paper is to examine the possible solutions towards the transition challenges, focusing on the situation in Africa which reveals the situation in most other developing nations or regions in the world. This will lead to a thorough look at the global experiment and awareness campaign on the world IPv6 day, which was setup to uncover the transition problems and develop strategies to resolve them.

The analysis carried out in this work will be limited to the African IPv6 deployment from 1984 to 2011 to justify the IPv6 promotion campaign realized through this 24-hour global experiment carried out every year.

Section two discussed the internet protocol (IP) versions; Internet Protocol version four (IPv4) and Internet Protocol version six (IPv6). The regional internet registries (RIRs) and their functions are also

presented. Section three discussed the transition from IPv4 to IPv6 the importance of the transition and the trend. Section four presents the results of the deployment statues of IPv6 in Africa, the situation in most of the countries in the region and the measure to improve the situation. Section five presents the conclusions and recommendations.

II. OVERVIEW

Internet Protocol is a set of technical rules that define how computers communicate over a network [6]. There are currently two versions [6]: IP version 4 (IPv4) and IP version 6 (IPv6). IPv4 was the first version of Internet Protocol to be widely used and still accounts for most of today's Internet traffic. There are just over 4 billion IPv4 addresses. While that is a lot of IP addresses, it is not enough to last forever. IPv6 is a newer numbering system to replace IPv4. It was deployed in 1999 and provides far more IP addresses [6], which are expected to meet the need in the future. All internet connected devices and websites have an IP address so that the internet's servers know where to send information to. When a website's address (or URL) is typed into a browser, the system needs to convert it into an IP address so that it knows which computer to connect to [9]. To do this, the system uses the internet's equivalent of a phonebook, known as the Domain Name System (DNS). At the moment, the vast majority of IP addresses in the DNS resolve to IPv4 - the current standard for addresses. So even if you have an IPv6-enabled machine that is connected to an IPv6-enabled network, you will still be connected to another computer or website using IPv4. Some websites have been set up to use IPv6, but generally you need to type in a special web address (such as <http://ipv6.sanger.ac.uk>, or <http://ipv6.google.com>) to connect using the new protocol [9].

A global experiment and awareness campaign on the World IPv6 day has been set up to uncover the transition problems and develop strategies to resolve them. Google, Facebook, Yahoo, Akamai, and Limelight Networks will be some of the organizations offering their content over IPv6 for a 24-hour 'test flight' [1] [14]. The goal is to motivate organizations-Internet-service providers, hardware makers, operating-system vendors and web companies-to prepare their services for IPv6, ensuring a successful transition as IPv4 addresses run out [14]. On World IPv6 day, the Sanger Institute, along with more than 300 organisations, advertise both IPv4 and IPv6 addresses in the DNS [9]. This will allow users with IPv6-enabled devices to connect via IPv6 without need to use the special address. IPv4 and IPv6 will coexist on the Internet for many years [7]. Users without IPv6 connectivity will continue to access the sites using IPv4 as normal [9] for the moment but with little or increasing restriction in the future. In comparison (Table 1) the major difference between IPv4 and IPv6 is the number of IP addresses. Although there are slightly more than 4 billion IPv4 addresses, there are more than 16 billion-billion IPv6 addresses [6].

Table 1: Comparing IPv6 and IPv4 [6]

	Internet Protocol version 4 (IPv4)	Internet Protocol version 6 (IPv6)
Deployed	1981	1999
Address size	32-bit number	128-bit number
Address format	Dotted decimal notation:192.168.0.202	Hexadecimal notation:3FFE:0400:2807:8AC9::/64
Number of addresses	2^{32}	2^{128}

2.1. Regional Internet Registries

Regional Internet Registries (RIRs) are independent, not-for-profit membership organizations that support the infrastructure of the Internet through technical coordination [2]. There are five RIRs in the world today (Figure 1). Currently, the Internet Assigned Numbers Association (IANA) allocates blocks of IP addresses and ASNs, known collectively as Internet number resources, to the RIRs, who then distribute them to their members within their own specific service regions. RIR members include Internet Service Providers (ISPs), telecommunications organizations, large corporations, governments, academic institutions, and industry stakeholders, including end users. The RIR model of open, transparent participation has proven successful at responding to the rapidly changing Internet environment. Each RIR holds one to two open meetings per year, as well as facilitating online

discussion by the community, to allow the open exchange of ideas from the technical community, the business sector, civil society, and government regulators.

The five RIRs are [2]:

AFRINIC - Africa region

APNIC - Asia and Pacific region

ARIN - Canada, many Caribbean and North Atlantic islands, and the United States

LACNIC - Latin America and parts of the Caribbean

RIPE NCC - Europe, Parts of Asia and the Middle East



Fig. 1: The RIRs and the general areas of responsibility (courtesy of NOR) [5]

Each RIR performs a range of critical functions including [2]:

- The reliable and stable allocation of Internet number resources (IPv4, IPv6 and AS Number resources)
- The responsible storage and maintenance of this registration data.
- The provision of an open, publicly accessible database where this data can be accessed.
- RIRs also provide a range of technical and coordination services for the Internet community.

2.2. IPv4 Current Status

The IPv4 address space is a 32 bit field. There are 4,294,967,296 unique values, considered in this context as a sequence of 256 "/8s", where each "/8" corresponds to 16,777,216 unique address values. In adding up these special purposes use address reservations there are the equivalent of 35.078 /8 address blocks in this category [11]. This is composed of 16 /8 blocks reserved for use in multicast scenarios, 16 /8 blocks reserved for some unspecified future use, 1 /8 block (0.0.0.0/8) for local identification, a single /8 block reserved for loopback (127.0.0.0/8), and a /8 block reserved for private use (10.0.0.0/8). Smaller address blocks are also reserved for other special uses. The remaining 220.922 /8 address blocks are available for use in the public IPv4 Internet. IANA holds a pool of unallocated addresses, while the remainder has already been allocated by IANA for further downstream assignment by the RIRs [11]. The current status of the total IPv4 address space is indicated in figure 2.

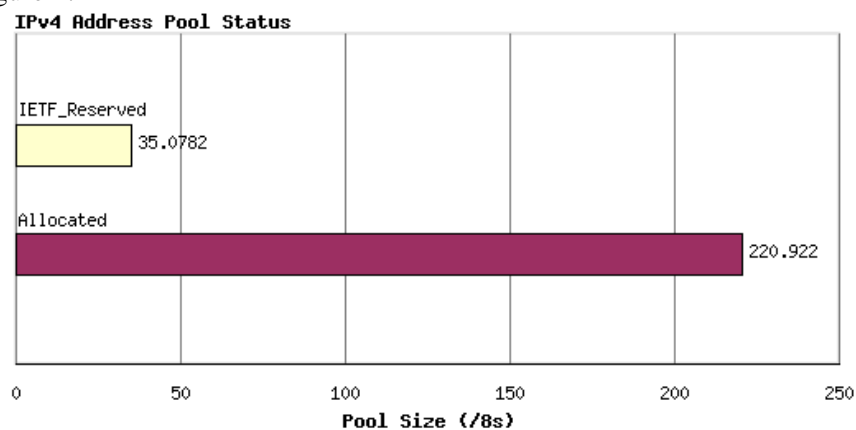


Fig. 2: Address Pool Status [11]

The current RIR address status (Table 2) which shows the present situation in the various RIRs based on the amount of the assigned addresses and the remaining addresses.

Table 2: The current RIRs address status

RIR	Assigned Addresses (/8s)	Remaining Addresses (/8s)
AFRINIC	8.5854	4.4107
APNIC	53.7944	1.2055
ARIN	77.9392	5.9865
LACNIC	15.3909	4.6091
RIPE NCC	45.2077	3.7923

Stephen Groat, et al [18] testified the eminence of the migration from IPv4 to IPv6 addresses. They argued that current methods of IPv6 address assignment, both stateful and stateless, use static identifiers that allow geotemporal tracking of users, monitoring and correlation of traffic over multiple sessions, and targeting for system-specific cyber-attacks. Susan Trulove [19] from Virginia Tech discussed the need for the IPv6 to replace the 20-year-old internet protocol version 4. Mark Tink [20] discussed the readiness for effects of IPv4 exhaustion, the dual-stack IP deployment on May 2007 and the transition from IPv4 to IPv6 addresses in Africa. John Loughney [21] carried out research on IPv4 allocation pool exhaustion and the switch to IPv6 addresses. He concluded that IPv4 addresses will run out, but there are going to be some dynamic issues which affect this. According to him, public IPv4 addresses may be needed for transition, so earlier usage of IPv6 can help. Silvia Hagen [22] in a brief study about IPv6 discovered that Asia especially Korea, China, Japan and India have embraced IPv6. She further argued that USA and Europe are planning for IPv6 deployment but nothing was found describing what Africa has done to prepare for IPv6 deployment.

III. THE TRANSITION

The internet is fast running out of addresses [9], [23]. By the end of the year it is thought that almost all the available addresses for computers, smartphones and websites will have been exhausted [9]. The best solution to ensure that the web can grow to its full potential is to change the way the system reads websites' addresses by moving to the next generation of addresses, known as IPv6. However, this system has never been used at a global scale and potential problems need to be uncovered before it can become the internet's new standard. The leading internet providers and websites collaborated in a 24-hour global experiment - World IPv6 day. The goal of this day is to tackle this pressing issue and drive forward the change needed ensure the continued delivery of full and free access to data and resources to the research community via the web [9].

On Wednesday 8 June, more than 300 organizations, institutions, companies and governments under went a trial experiment of the future language of the internet: World IPv6 day. From 1am Nigerian time to 1am (Nigeria time) on Thursday morning, alongside Facebook, Google, Cisco, Yahoo, Sony, universities and many US Government departments, some institutions such as the Sanger Institute opened its websites to visitors using two methods of delivery [9]: the current standard of IPv4 and the future standard of IPv6.

This change is needed because IPv4 is about to run out of addresses [23] for all the websites, computers, smartphones and internet-enabled devices that are coming on stream. In fact, the last major batch of available IPv4 addresses (about 80 million of them) was given out in a ceremony on 3 February 2011 [9], [15] [24]. It is expected that all these addresses will have been taken by September 2011 [9].

The move to IPv6 is facing challenges. Although the new addressing system was designed in the 1990s and its technical foundations are now well established, not everyone is using currently equipment that can handle IPv6. It is estimated that roughly 0.05% of internet users will not be able to connect to websites that offer both IPv6 and IPv4 (a system known as 'dual stacking') [9]. IPv4 and IPv6 will coexist on the Internet for decades, creating the need for additional transition mechanisms because the dual-stack model won't solve all of the emerging problems for network operators [25].

To uncover any problems that might occur with a switch to IPv6, and to create an event to drive forward the change, the Internet Society is coordinating a global experiment. The society - a charity dedicated to ensuring the open development, evolution and use of the internet - is using this day-long

event, to create a common focus to bring together all stakeholders in the internet to resolve any issues [9]: from governments and universities, to internet service providers, operating system suppliers and hardware manufacturers.

3.1. IPv6 Deployment Trend

Globally, the rate of the IPv6 deployment is growing super-linearly [17] (Figure 3) showing the response towards the adoption of the future promising technology.

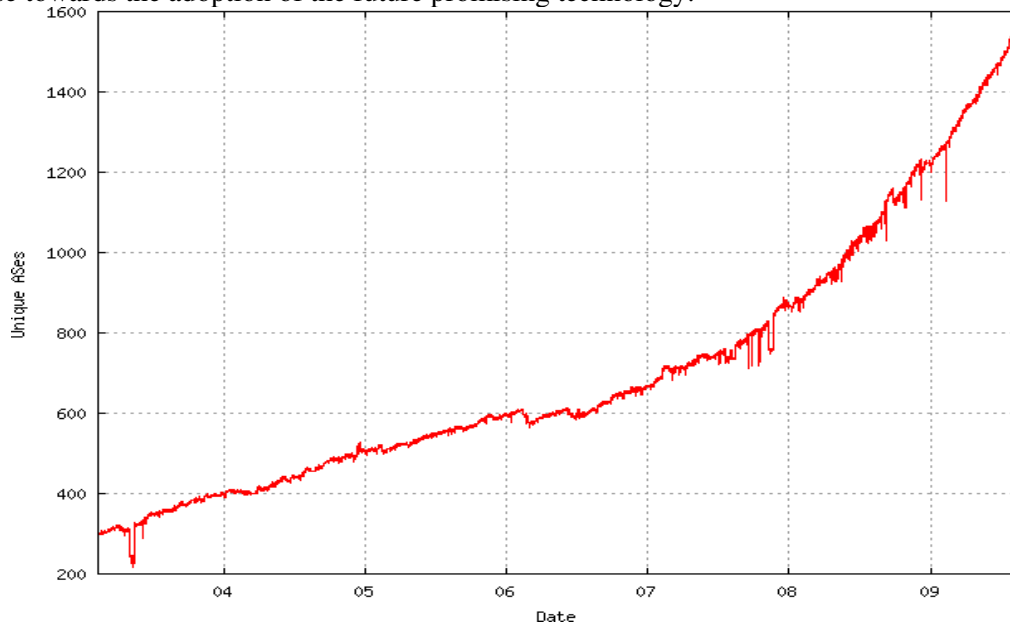


Fig. 3: IPv6 deployment growth [17]

The NRO announced that the rate of new entrants into the IPv6 routing system has increased by approximately 300 percent over the past two years. This growth has been attributed to the active promotion of IPv6 by the five RIRs and their communities [2].

IV. RESULTS AND DISCUSSION

The IPv6 adoption and deployment status in Africa was drawn from 1984 to 2011 as shown in figure 4. The results were based on the latest information on IPv6 Addresses allocated in the AfriNIC region [4]. The deployment which was further categorized into countries (Figure 5) in the region also revealed the poor rate of IPv6 deployment, recorded in most the countries.

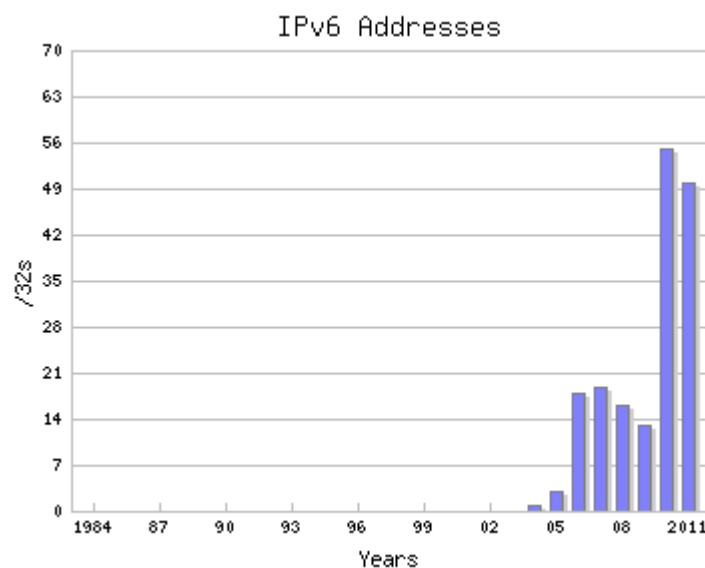


Fig.4: IPv6 Addresses (Yearly) [4]

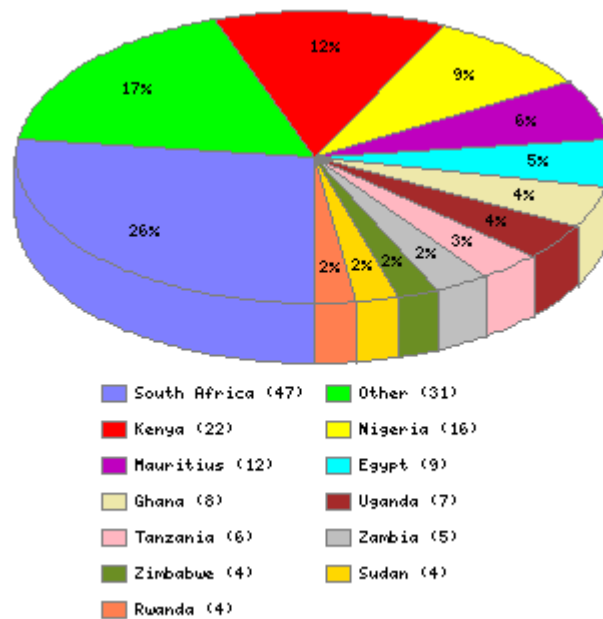


Fig.5: IPv6 Address by Country [4]

Note that “Other” refers to all countries that have less than 2% of the total.

South Africa recorded the highest percentage of the IPv6 address allocation followed by Kenya and Nigeria with 26%, 12% and 9% respectively (Figure 5). According to the result, twelve countries in Africa recorded 2% and above of the IPv6 address allocation and the remaining greater number of countries recorded less than 2% of the allocation indicating little or no awareness of the future internet communication technology in most countries in Africa.

The poor IPv6 adoption in Africa as reflected in the results was attributed to the following:

- People are not aware of additional benefits IPv6 should bring in their lives; thereby, nothing is motivating them especially in the dynamic sustenance towards IPv6. The main cause is the fact that there is no active e-strategy program in the different government/ political policy agenda.
- The continent on its entirety lacks the general IPv6 consciousness-raising campaign for total capacity building programs. Many organizations including AFRINIC tried to initiate some capacity building programs upon IPv6 for the attention of mainly some technical bodies such as the Internet Service Providers (ISPs). Since then, not sure those things are going ahead at the expected speed. It should not be strategic to begin the consciousness-raising campaign focusing on those specific bodies.
- Many ISPs in Africa are not yet aware on the technical/business benefits of IPv6. Some of the ISPs in Africans are still relying on the simple seem fact that since IPv4 is working well presently, it is not cost-effective should they change or remove it.

4.1. The Consequences of Delay in Ipv6 Deployment

Without a dual-stacked network or deployed protocol translation services, an individual user gaining Internet access for the first time from an IPv6-only ISP may not be able to access the Web sites or mail servers for organizations that operate IPv4-only networks [8]. There are implications to not adopting IPv6. These implications become evident as wide-scale deployment of IPv6 accelerates. Not adopting IPv6 may cause the following types of issues for the various types of Internet users [8]:

Individual Users: Individual users may not be able to view Web sites and communicate with certain destinations. Many individuals use the Internet to communicate with distant friends and family, research medical issues, and participate in group discussions among other things.

Enterprise Organizations: Enterprise organizations and corporations may not be able to communicate with certain critical government resources, clients, and potential customers. E-mail is a critical form of communication for most enterprise organizations today and their Web sites are vitally important resources for them to communicate with the public.

Governments: Governments may lose their ability to see and communicate with the “whole Internet.” Access to information is critical for governments. There also may be an inability for citizens and other Internet users to access information about the government and communicate over the Internet with government agencies.

Service Providers: Organizations that provide services over the Internet may experience customer and/or revenue losses if they do not update their offerings to include IPv6. Customers will expect to be able to communicate with everyone else on the Internet and may seek out other ways to do that if their current service provider is not capable.

4.2. The Way-out

In order to facilitate the rate of IPv6 deployment, the various levels of governance (government, academic/ business organizations, etc) in African countries should in their day-to-day policy making include as an agenda, active e-strategy programs to help not only in the imperative deployment of IPv6 but also in other Information and Communication Technology (ICT) facility development. AFRINIC and other related organizations in Africa should be expected to develop some capacity building programs upon IPv6 awareness campaign for the general populace, thereby motivating the people and enlightening them about the technical/ business benefits derived from switching to IPv6.

V. CONCLUSION AND RECOMMENDATION

The Internet Protocol Version 6 (IPv6) had been adopted globally to solve the future problem of Internet Protocol Version 4 (IPv4) exhaustion. The IPv4 was the first Internet protocol version to be used widely and still dominating today's internet traffic. The global rate of IPv6 deployment accelerated due to the total adoption of the technology in most developed regions of the world. The African situation was revealed to be different as reflected by the poor results of IPv6 deployment rate recorded in most countries of the region by AfriNIC over some past years.

The various development strategies stipulated to aid the deployment of IPv6 such as the global awareness campaign of the World IPv6 day, general capacity building set up by the RIRs, and many more, have been confirmed effective with the recent IPv6 deployment improvement recorded globally even in Africa.

It was recommended that the business sector should start to support IPv6 by hosting content on IPv6-enabled websites, ensuring accessibility to IPv6 users. Software and hardware vendors should implement IPv6 support in their products urgently, to ensure they are available at production standard when needed. Governments should learn more about IPv6 transition issues in order to support IPv6 deployment efforts in their countries. IPv6 requirements in government procurement policies are critical at this time. Finally, civil society, including organizations and end users, should request IPv6 services from their ISPs and vendors, to build demand and ensure competitive availability of IPv6 services in coming years.

REFERENCES

- [1]. Carolyn, D. M., (2011). *World IPv6 Day: Tech Industry's most-watched event since Y2K* <http://www.networkworld.com/news/2011/060711-ipv6-expect.html>
- [2]. Paul, W., (2008). *IPv6 Growth Increases 300 Per Cent in Two Years*. <http://www.ripe.net/internet-coordination/news/industry-developments/ipv6-growth-increases-300-per-cent-in-two-years>
- [3]. Philemon, (2007). *Africa and IPv6*. <https://www1.ietf.org/mailman/listinfo/ietf>
- [4]. AfriNIC, (2011). *Statistics – IPv6 Recourses*. http://www.afrinic.net/statistics/ipv6_resources.htm
- [5]. Michael, K., (2008). *IPv6: What you need to know*. <http://techrepublic.com.com>
- [6]. Arin, (2008). *simple overview of IPv6 and the differences between it and IPv4* <http://techrepublic.com.com>
- [7]. Arin, (2008). *IPv4 and IPv6 coexistence — what does that mean?* <http://techrepublic.com.com>
- [8]. Arin, (2008). *What really happens to my company Internet access if it or my ISP network doesn't transition in time?* <http://techrepublic.com.com>
- [9]. Phil, B., (2011). *24 hours to shape the Internet's future* http://www.sanger.ac.uk/about/press/features/ipv6day.html#t_pg3
- [10]. Geoff, H., (2003). *The Myth of IPv6*. The Internet Protocol Journal, Volume 6, No. 2.
- [11]. Geoff, H., (2011). *Current Status*. <http://www.potaroo.net/tools/ipv4/#r4>

- [12]. Takashi, A., (2011). *IPv4 Address Report 2011; Projected RIR Unallocated Address Pool Exhaustion*. <http://inetcore.com/project/ipv4ec>
- [13]. Adiel, A., Alian, P., (2010). *Analysis of the future exhaustion of the IPv4 Central pool in relation to IANA and its impact on the AfriNIC region*. http://www.afrinic.net/news/ipv4_exhaustion.htm
- [14]. Michael, K., (2011). *How new research aims to protect our privacy on IPv6 networks*. <http://www.techrepublic.com/blog/security/how-new-research-aims-to-protect-our-privacy-on-ipv6-networks/5583?tag=nl.e036>
- [15]. Matthew, F., (2011). *Edging Toward the End of IPv4: A New Milestone in the History of the Internet*. <http://iso.org/wp/ietfjournal/files/2011/03/IETFJournal79FINAL.pdf>
- [16]. Geoff, H., (2003). *IPv4: How long do we have?* The Internet Protocol Journal, Volume 6, No. 4
- [17]. Derek, M., (2011). *What is the Status of IPv6 Outside of Penn State?* <https://wikispaces.psu.edu/display/ipv6/IPv6+FAQs>
- [18]. Stephen, G., (2009). *Dynamic Observation of IPv6 Addresses to Achieve a Moving Target Defense*. Virginia Tech Intellectual Properties.
- [19]. Susan, T., (2011). *Greatest Winning Network Security*. Blackburg ,Virginia Tech.
- [20]. Mark, T. *IPv6 Deployment Africa Online Zimbabwe*. Global Transit, Kuala Lumpur, Malaysia.
- [21]. John, L., (2008). *Converged Communication and IPv6*. Nokia.
- [22]. Silvia, H. *IPv6 Deployment in Africa*. <http://searchnetworking.techtarget.com/answer/IPv6-deployment-in-Africa>
- [23]. Christian, J., (2011). *Zero Address, One Solution Two Problems*. <http://iso.org/wp/ietfjournal/?p=2187#more-2187>
- [24]. Olaf, K., (2011). *Words from the IAB Chair*. <http://iso.org/wp/ietfjournal/?p=2162#more-21862>
- [25]. Carolyn, D.M., (2011). *IPv4, IPv6 Coexistence Changes Network Operators*. <http://iso.org/wp/ietfjournal/?p=2173#more-2173>

Authors

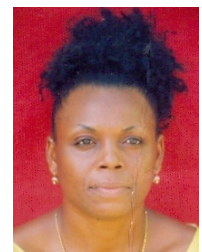
Emmanuel Chukwudi Agbaraji completed his Bachelor of Engineering (B.ENG.) degree in Electrical Electronic Engineering in 2006 from Federal University of Technology Owerri (FUTO). He attained Microsoft Certified Professional (MCP) in 2010. His research interest includes software engineering, data communication and internet computing. At present Mr. Agbaraji is working on his thesis for the award of Masters of Engineering in computer engineering.



Felix Kelechi Opara is presently the Head of Electrical Electronic Engineering department in Federal University of Technology Owerri (FUTO). He holds a PhD from FUTO, where he completed his Masters of Science and Bachelor of Engineering degrees. Engr. Dr. Opara has attained engineering certified professional. His research interest includes data communication, software engineering, and protocol development.



Arvis Ijeaku Aririguzo is a lecturer in Federal Polytechnic Nekede. She holds a Bachelor of Engineering (B.ENG) degree in Electrical Electronic Engineering from Federal University of Technology Owerri. She has attained engineering certified professional and Cisco Certified Network Associate (CCNA). Her research interest includes data communication and internet computing. Engr. Mrs. Aririguzo is presently working on her thesis for the award of Masters of Engineering in computer engineering.



STUDY AND REALIZATION OF DEFECTED GROUND STRUCTURES IN THE PERSPECTIVE OF MICROSTRIP FILTERS AND OPTIMIZATION THROUGH ANN

Bhabani Sankar Nayak¹, Subhendu Sekhar Behera², Atul Shah¹

¹ BTech, Department of ECE, National Institute of Science & Technology, Berhampur

² BTech, Department of CSE, National Institute of Science & Technology, Berhampur

ABSTRACT

Defected ground structures (DGS) have been developed to enhance different characteristics of many microwave devices. In this paper a Micro-strip low pass filter with Dumbbell Shaped Slot Defected Ground Structure (DGS) is designed. The response of the filter is analyzed with respect to variation in dimension of the DGS unit. The variation of dimensions of defects studied with their corresponding change in capacitance, inductance as well as frequency response. The defects dimensions are modeled with respect to frequency using the artificial neural network. Optimizing the convergence of Artificial Neural Network (ANN) classifiers is an important task to increase the speed and accuracy in the decision-making. The frequency response of the micro strip filter is modeled with respect to the variation in dimension of DGS using CST microwave studio. The dimensions are further optimized in order to achieve minimum error in frequency response. Incremental and online back propagation learning approach is followed in the training of neural network because of its learning mechanism based upon the calculated error and its ability to keep track of previous learning iteratively. The simulation results are compared with the results obtained through ANN and the designs are further optimized.

KEYWORDS: Filters, defected ground structures, ANN, CST microwave studio.

I. INTRODUCTION

Defected Ground Structures (DGS) have been developed in order to improve characteristics of many microwave devices [1]. Most of its advantages lies in the area of the microwave filter design, microwave oscillators, microwave couplers as well as microwave amplifiers. DGS is motivated by the study of Electromagnetic band gap structures [2]. It is more easily an LC equivalent circuit. Presently there are vast applications of microwave components such as filters, amplifiers, couplers, antennas in various fields like mobile radio, wireless communication, and microwave millimeter wave communication [4]. Basically micro strip technology consists of transmission line made of conducting material on one side of dielectric substrate with the ground plane on other side. A microwave filter is a two- port network used to control the frequency response at a certain point in a microwave system by providing transmission at frequencies within the pass band of the filter and attenuation in the stop band of the filter. Defected ground structures (DGS) are recently one of the hottest topics which are researched in microwave domain, which developed from the photonics band gap (PBG) structures [1]. The high characteristic impedance of DGS is also used in digital systems [2]. DGS is an etched lattice which makes one or a few of PBG etched ground elements in the ground plane. DGS can achieve high-performance which can not be obtained by conventional technology. Because of the advantage of DGS, such as having small structure size, more transitional sharpness, achieving broader stop band responses, high characteristics impedance and simple model, it has been widely used in the design of microwave filter. The defects in the ground plane of the transmission lines [3] such as dumbbell, elliptical, square etc disturbs the shield current distribution and also changes the characteristics of

transmission lines e.g. capacitance and inductance. The series inductance due to the DGS section increases the reactance of a micro strip with the increase of the frequency. Thus, the rejection of the certain frequency range can be started. The parallel capacitance with the series inductance provides the attenuation pole location, which is the resonance frequency of the parallel LC resonator. However, as the operating frequency increases, the reactance of the capacitance decreases. Thus, the band gap between the propagating frequency bands can be occurred. By etching DGS on the ground plane it is possible for the designer to increase the equivalent inductance L highly and to decrease the equivalent capacitance C at the same time, and finally to raise the impedance of the micro strip line to a level more than 200Ω [3]. But the problem arises, as there is no fixed mathematical model in order to relate the frequency response with respect to the change in dimension of DGS Unit cell. Our main focus lies in optimizing the frequency response with the help of ANN being trained with Back Propagation Algorithm [14]. Back Propagation is the most popular neural network training algorithm for supervised learning with weight correction rule [11]. The weight correction methodology comprises of back-propagating the errors from output layer to hidden layer, thus finding the optimal set of weights. It is used in a greater extent in the field of Data Analysis, Weather Forecasting, and Trading Analysis etc. As the Learning Rate has a significant effect on the results, we choose the best through iteration. This allows the Back Propagation to be optimized. The design procedure is presented in the section.2 along with its response due to the variation of different dimensions of DGS. The designs are implemented using CST microwave studio and the results are analyzed. In the 3rd section we implemented the back propagation neural network to model the frequency response with respect to the dimension of DGS. The application of artificial neural network ensures an optimum design methodology for microstrip filter design which is revealed when comparing the results with analytical methods and the results of the simulation software's [14]. The designs are made using the CST microwave studio software [15] and also the simulations for analyzing the frequency response for every change in dimensions of DGS, calculation of inductance and capacitance etc. ANN algorithm is implemented using C programming in DEV C++ compiler and the results obtained for training and testing is plotted with the help of MATLAB [15].

II. A STUDY ON RELATED WORK

There has been a lot of research on optimization of frequency response using different soft computing algorithms. A novel approach for calculating the resonating frequency of microstrip antenna is presented in [14] by R.K. Mishra and A. Pattnaik. In the reference [4] A part of optimization is made to model the frequency response of the planar microstrip antenna with respect to the change in dimension of DGS. There are several algorithms to optimize the training process. Back Propagation is one of the most popular neural network training algorithms for supervised learning. The weight corrections are updated with a generalized delta rule to minimize the prediction error through iterations. There have been similar attempts made to choose the dielectric constant for antenna design using Neural network model [11]. In reference [13] a new methodology for determining the input impedance for microstrip antenna is presented. In this paper we have implemented the Artificial neural network algorithm to model the frequency response of microstrip filter with respect to the dimensions of dumbbell shaped DGS. The weight correction methodology comprises of back-propagating the errors from output layer to hidden layer, thus finding the optimal set of weights [7-9]. In this paper a feed-forward network (FFN) has been considered. FFN allows the signal to flow from input to output layer in feed-forward direction [7, 9].

III. DESIGN OF FILTER AND RESPONSE DUE TO DEFECTED GROUND

The low pass filter configuration having five sections of alternating high and low impedances is shown in the figure1. The lpf was designed using the formulations depicted in [3].The order of filter designed is of 5th order. The Dumbbell Shaped Slot DGS section is fully described by two parameters the etched lattice dimension and gap distance. The influences of these two parameters on frequency characteristics of a micro strip are shown by simulations. All simulations were carried out on CST Microwave studio. The dimension of DGS slot are given below in fig2 as l, w, g respectively.

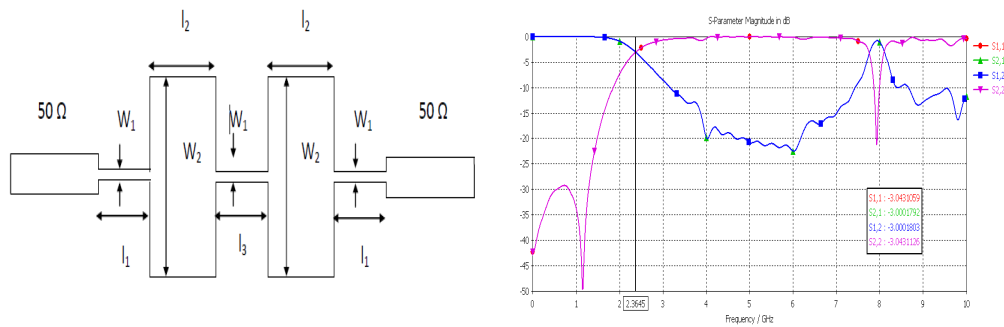


Fig1.design of low pass filter and its response

(Where the dimension are given by $w_1 = 0.293$ mm, $w_2 = 6.352$ mm, $l_1 = 2.917$ mm, $l_2 = 7.1323$ mm, $l_3 = 11.036$ mm, and the corresponding inductance and capacitance are given as $L_1 = 2.05$ nH, $C_2 = 2.1472$ pF, $L_3 = 6.634$ nH, $C_4 = 2.146$ pF, $L = 2.05$ nH)

When the single dumbbell shaped slot is placed at the center, it provides inductance and hence by placing the DGS in the structure, effective inductance increases and the cut off frequency decreases.

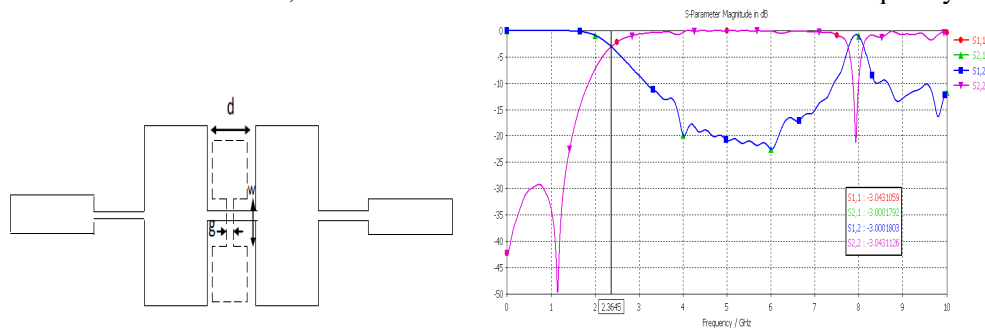


Fig2.design of low pass filter with defects and its response

The line width is chosen to be the characteristic impedance of 50Ω micro strip line for simulations. Three DGS unit circuits were simulated with the different dimensions. In order to investigate the influence of the square lattice dimension, the etched gap, which is related with the gap capacitance, was kept constant to 0.1 mm for all three cases and the etched square area was varied. The substrate with 0.762 mm thick and a dielectric constant of 3.2 is used for all simulations. We observe that employing the proposed etched lattice increases the series inductance to the micro strip line. This effective series inductance introduces the cutoff characteristic at certain frequency. As the etched area of the unit lattice is increased, the effective series inductance increases, and increasing the series inductance gives rise to a lower cutoff frequency, as seen in Table 1. There are attenuation poles in simulation results on the etched unit lattices. These attenuation poles can be explained by parallel capacitance with the series inductance. This capacitance depends on the etched gap below the conductor line [4]. The capacitance values are identical for all cases due to the identical gap distance. However, the attenuation pole location, which corresponds to the resonance frequency of the parallel LC circuit, also becomes lower because as the series inductance increases, the resonance frequency of the equivalent parallel LC circuit decreases. The results are shown in table 1.

Table1 variation of length and gap in DGS

Variable(unit)	d =7	d =8	d =9
Inductance(nH)	5.24	6.39	7.56
Capacitance(pF)	0.70	0.69	0.67
Cut off freq (GHz)	1.70	1.48	1.34
Center freq (GHz)	2.59	2.36	2.21
	G=0.1	G=1	G=2
Inductance(nH)	3.42	3.58	3.70

Capacitance(pF)	0.72	0.18	0.08
Cut off freq (GHz)	2.25	3.4	3.52
Center freq (GHz)	3.16	7.14	8.5

The lattice dimension is kept constant to 5 mm for all three cases and the etched gap distance is varied. Due to the constant lattice dimensions, we can expect that the effective series inductances are also constant for all cases. There is no change in cutoff frequency despite the variation of the gap distance. This means that the gap distance does not affect the effective series inductance of a micro strip. Variation of the effective capacitance only affects the attenuation pole location[1]. As the etched gap distance increases, the effective capacitance decreases so that the attenuation pole location moves up to higher frequency. When the single dumbbell shaped slot is placed at the center, it provides inductance and hence by placing the DGS in the structure, effective inductance increases and the cut off frequency decreases. When the single dumbbell shaped slot is placed at the center, it provide inductance and hence by placing the DGS in the structure, effective inductance increases and the cut off frequency decreases. Response is improved in terms of sharpness because of decrease in the capacitance. The Cut off frequency of the low pass filter is 1.66 GHz and the slope is 9.65 dB/GHz. When g is reduced to 0.1 mm the effective capacitance increases which results in lowering of attenuation pole location. The insertion loss reaches -50 dB down. As the area of the slot is kept constant, there is no change in effective inductance and hence the cut off frequency is constant. When the width of the etched slot is decreased effective inductance is decreased because of which cut off frequency is increased. Also the response is improved in terms of insertion loss and return loss.

Table2 . Variation with respect to the change in d

S. No	D(mm)	Cutoff frequency(GHZ)	Slope (dB/GHz)
1	6.3	2.4214	7.4808
2	6.1	2.4434	7.3361
3	5.9	2.4787	7.13

According to the Quasistatic Theory of DGS depicted in [4] the electric and magnetic fields are mostly confined under the microstrip line. The return current on the ground plane is the mirror image of the current distribution occurred at the strip line. The maximum surface current lies over the ground plane and the width of side filament arm which contribute to the inductance of DGS [4]. The gap is represented by the equivalent capacitances, the inductances and capacitances are derived from the physical dimensions using quasi-static expressions for microstrip crosses, lines and gaps given in [5]. The electrical equivalent model of DGS is given below [4,6] .it is been observed that for various change in dimension of DGS we are getting a different frequency for which any mathematical model is not established yet. So for the simplification we are implementing neural network in order to model the frequency change and optimize the design. In the next section we implemented the artificial neural network using Dev CPP compiler [15] for the training and testing of the network. This is validated with the simulations made at CST microwave studio on the desired set of testing data sets and also the frequency response is checked.

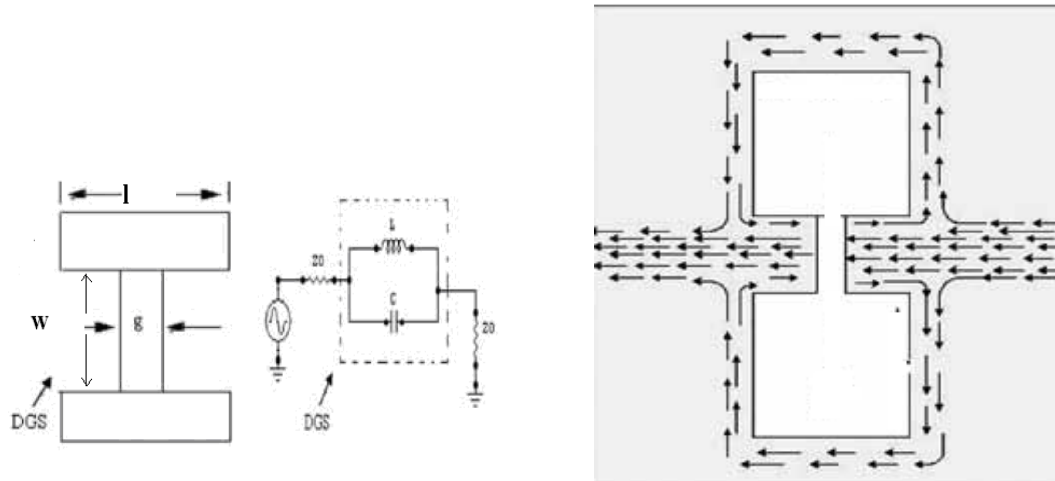


Fig 3. Equivalent circuit of DGS

IV. OPTIMIZATION THROUGH ANN

Artificial neural network has been implemented to determine the problem of accurate determination of frequency of dumbbell shaped DGS for a desired dimension of DGS. The input to the ANN model are the dimension of defects l, w, g and the target data is frequency.

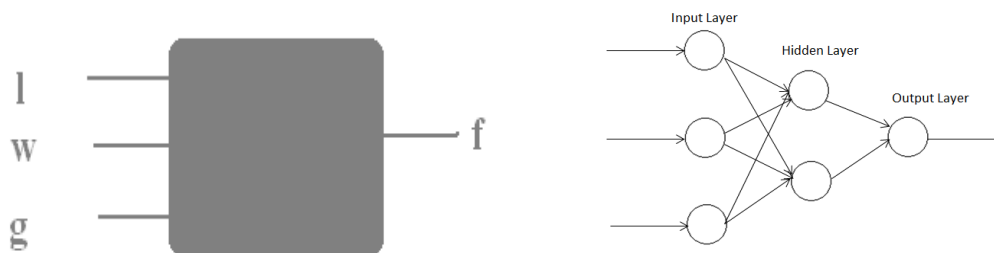


Fig 4 . ANN model of DGS

Back Propagation algorithm is implemented which comprises of two phases. First, a training input pattern is presented to the network input layer which is propagated forward to the output layer through hidden layers to generate the output. If this output differs from the target output presented then an error is calculated (Here Mean Square Error). This error is back-propagated through the network from the output layer to the input layer and weights are updated [7].

As we are not satisfied with a normal back propagation, we investigated the results with learning rate starting from 0.1 to 1.0 with momentum constant equal to 0.9 to speed up the learning process. The epoch size for each learning rate is 20 epochs [7,8]. The cost function used here is the Mean Square Error (MSE). The log sigmoid function in equation 1 is used as the transfer function associated with the neurons in hidden and output layer to obtain the normalized [0, 1] nodal outputs.

$$f(x) = (1 + e^{-x})^{-1} \quad (1)$$

As we use log sigmoid as the transfer function, we normalize the input values [0,1]. It will reduce calculation complexities. The class values for each dataset are also normalized in the range from 0 to 1.

V. ALGORITHM

Step1: Set Learning Rate $\lambda = 0.1$, Momentum Constant $\alpha = 0.9$. Initialize No. of Tuples according to dataset. Initialize set of weights randomly.

Step2: Set $MSE_{total} = 0$ and $i=0$.

Step3: Present i^{th} input vector $X_{i0}, X_{i1}, \dots, X_{iN-1}$ and specify the desired output d_{i0} . Calculate actual output Y_{i0} and MSE_i .

Step4: Modify the weights starting from output layer to input layer using delta rule given below.

$$W_{jk}(t+1) = W_{jk}(t) + \lambda \delta_k x_j' + \alpha (W_{jk}(t) - W_{jk}(t-1)) \quad [2]$$

Where $W_{jk}(t)$ is the weight from node j to node k at time t ; α is momentum constant; x_j' is either the output of node j or is input j ; λ is learning rate; and δ_k is an error term for node k . If node k is an output node, then

$$\delta_k = y_k (1 - y_k) (d_k - y_k) \quad [3]$$

Where d_k is the desired output of node k and y_k is the actual output.

If node k is an internal hidden node, then

$$\delta_k = x_j' (1 - x_j') \sum_l \delta_l W_{kl} \quad [4]$$

Where l is over all nodes in the layer above node k .

Step5: $MSE_{total} = MSE_{total} + MSE_i$.

Step6: Repeat by going to step3 if $i < \text{No. of Tuples}$.

Step7: $MSE_{total} = MSE_{total} / \text{No. of Tuples}$. Store MSE_{total} .

Step8: Repeat Step 2-6 for no. of epoch size.

Step9: $\lambda = \lambda + 0.1$. Repeat by going to step2 with initialization of weights randomly if $\lambda \leq 1.0$.

The ANN model is shown above with dimensions 1, g, w and cut off frequency obtained from the output of ANN for the chosen dielectric substrate. The design is made and parametric variations are obtained for 80 observations, 60 are used for training and rest 20 is used for testing. The best learning rate is chosen by testing each starting from 0.1 to 1.0. The learning rate chosen at 0.1 turned to be the best learning rate. The neural network with 2 neurons in 1 hidden layer and best learning rate reduces the error to 0.003401 in 20 epochs only while testing the neural network. The obtained results from ANN were checked by designing in CST and the frequency response were matched mostly with respect to the result obtained in ANN. Incremental back propagation learning approach is followed in the training of neural network[9,10]. The training result is shown below with the least error of 0.003401.

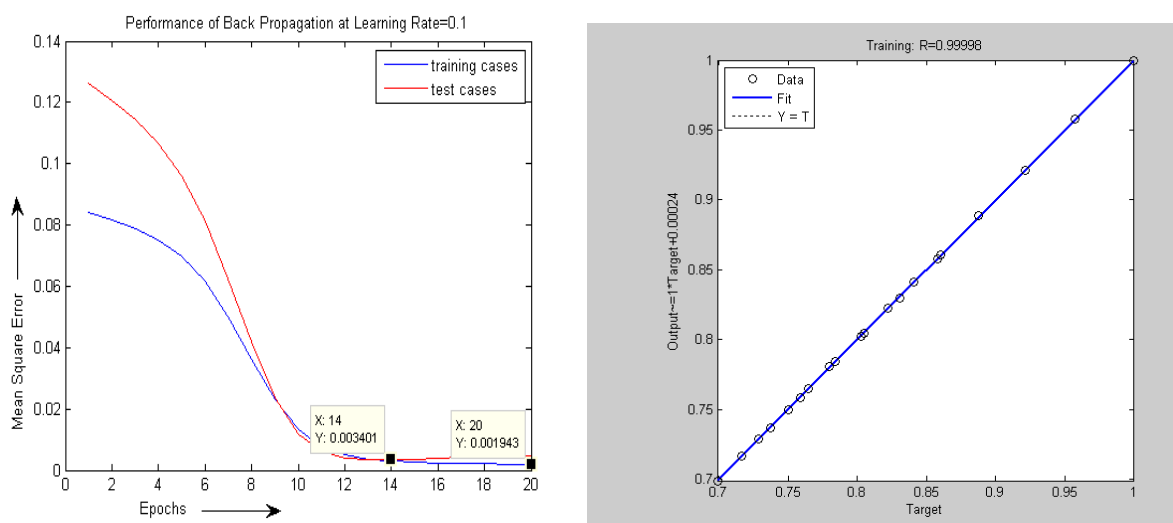


Fig 5 ANN result and regression plot

After the neural network model is created different values of l , w , g are taken and the frequency response is calculated with the help of artificial neural network and the results are cross checked with the help of CST microwave studio, it is observed that neural network works efficiently in determining the accurate frequency of the microwave filter with dumbbell shaped DGS. When we chose $l = 10$, $g = 0.1$, $w = 5$ (in mm), the neural network output was found out to be 1.649 GHz, where the simulated result is shown below which shows the frequency response at 1.6527 GHz. The response is shown below in fig 6 which shows that neural network is working efficiently with least MSE 0.003401

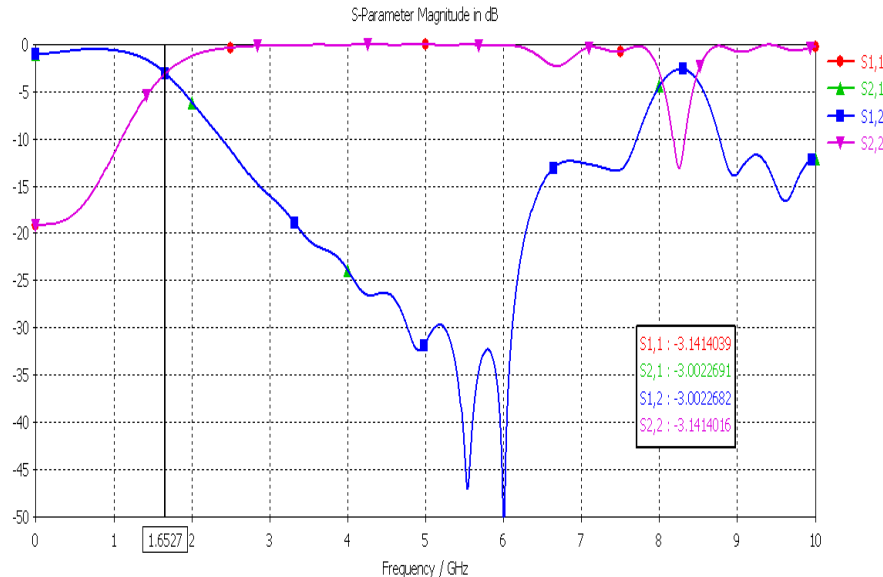


Fig 6 frequency response obtained with CST

VI. CONCLUSION

We designed the five elements LPF. After that dumbbell shaped defect is created in the ground plane. The addition of defects enhances the response of the filter as well as reduces the size. It consists of L-C parallel circuit having a resonant frequency characteristic. It is having band gap property, which is used in the many microwave applications. The frequency response of the dumbbell shaped defect is studied with respect to the dimension of its length, width and gap. The applications of artificial neural network for getting the frequency response of filter with respect to the dimensions of defects are done with the minimal error of 0.003401. Although training ANN model has spent little extra time, the successful intelligent model can quickly provide precise answers to the task in their training values range. Neural network efficiently worked to model frequency and the dimensions are optimized to give rise better response with the least error of 0.003401. The learning rate is chosen to be highly optimized one through iterations. The future scope of the work lies to implement Adaptive neuro fuzzy inference system (ANFIS) for the optimization and modeling of frequency response of microwave circuits, which will have better learning approach and higher degree of accuracy at a shorter time in comparison to ANN shown in [16].

ACKNOWLEDGEMENTS

The authors would like to thank, CST Company, India for their support in CST EM tool. We are grateful to Prof. Rabindra Kishore Mishra, for his kind suggestions and guidance in this work. We are thankful to P.K. Patra and A.K. Panda for their kind help. The authors are grateful to the anonymous reviewers for their constructive & helpful comments & suggestions.

REFERENCES

- [1] Kim, J. P. and Park, W.S., "Microstrip lowpass filter with multislots on ground plane," Electronics Letters, Vol. 37, No. 25, pp. 1525–1526, Dec. 2001.
- [2] Yang, F., and Rahmat Samii, Y., "Electromagnetic band gap structures in antenna engineering", Cambridge University press, USA, 2009

- [3] Bharathi Bhat , Shibani K. Koul “Stripline-like Transmission Lines for Microwave Integrated Circuits” *New Age International (p) Ltd, Publishers*,2007..
- [4] A.K Arya,M.V kartikeyan ,A.pattnaik “Defected ground structures in microstrip antenna: a review” in *frequenz* 64(2010)
- [5] Roy, S.M., Karmakar, N.C., Balbin, I., “Quasi-Static Modeling of Defected Ground Structure”, *IEEE Transaction on Microwave Theory and Technique*, vol.54, no.5, pp.2160 –2168, may 2006..
- [6] Oskouei, H.D., Forooghi, K. and Hakkak, M., “Guided and leaky wave characteristics of periodic defected ground structures”, *Progress in Electromagnetics Research, PIER* 73, 15 –27, 2007
- [7] Jung I. and Wang G., “Pattern Classification of Back-Propagation Algorithm Using Exclusive Connecting Network”, *World Academy of Science, Engineering and Technology* 36 2007.
- [8] Chattopadhyay S., Kaur P., Rabhi F., and Acharya R.U., “An automated system to diagnose the severity of adult depression”, *EAIT*, 2011
- [9] Jha G.K., “Artificial Neural Network”, *IARI*, 2006.
- [10] http://www.softcomputing.net/ann_chapter.pdf
- [11] Patnaik, A., Mishra, R.K., Patra, G.K., Dash, S.K., “An artificial neural network model for effective dielectric constant of microstrip line,” *IEEE Trans. on Antennas Propagation*, Vol. 45, no. 11, p. 1697, November 1997.
- [12] S.S pattnaik , D.C panda, S devi “A novel method of using Artificial Neural Networks to calculate input impedance of circular microstrip antenna”
- [13] Bailer-Jones D.M.and Bailer-Jones C.A.L., “Modelling Data: Analogies in neural networks, simulated annealing and genetic algorithms” in *Model-Based Reasoning: Science, Technology, Values, L. Magnani and N. Nersessian (eds.)*, New York: Kluwer Academic/Plenum Publishers, 2002.
- [14] R.K.Mishra and Patnaik, “*Designing Rectangular Patch Antenna Using the Neurospectral Method*”, *IEEE Transactions on Antennas and Propagation*,AP-51,8 August 2003,pp.1914-1921.
- [15] CST Design Studio , MATLAB and DEV CPP
- [16] Guney, K. and N. Sarikaya, “A hybrid method based on combining artificial neural network and fuzzy inference system for simultaneous computation of resonant frequencies of rectangular, circular, and triangular microstrip antennas,” *IEEE Trans. Antennas Propagat.*, Vol. 55, No. 3, 659–668, 2007.

Biographies

Bhabani Sankar Nayak is currently pursuing his B.Tech in dept of Electronics and communication engineering at National Institute of Science & Technology, Berhampur, Orissa. His research interest include Electromagnetic, antenna, microwave circuits, CAD, soft computing. He is currently working as research scholar at NIST under scholarship program.



Subhendu Sekhar Behera is currently pursuing his B.Tech in dept of computer science & engineering at National Institute of Science & Technology, Berhampur, Orissa. His research interest include soft computing, web designing, algorithm design.



Atul Shah is currently pursuing his B.Tech in dept of Electronics and communication engineering at National Institute of Science & Technology, Berhampur, Orissa. His research interest include intelligence system design, embedded systems.



ANALYSIS OF DISCRETE & SPACE VECTOR PWM CONTROLLED HYBRID ACTIVE FILTERS FOR POWER QUALITY ENHANCEMENT

Jarupula Somlal¹, Venu Gopala Rao Mannam²,

¹Assistant Professor, Department of EEE, K L University, Vijayawada, A.P, India

²Professor & Head, Department of EEE, K L University, Vijayawada, A.P, India

ABSTRACT

It is known from the fact that Harmonic Distortion is one of the main power quality problems frequently encountered by the utilities. The harmonic problems in the power supply are caused by the non-linear characteristic based loads. The presence of harmonics leads to transformer heating, electromagnetic interference and solid state device mal-functioning. Hence keeping in view of the above concern, research has been carried out to mitigate harmonics. This paper presents an analysis and control methods for hybrid active power filter using Discrete Pulse Width Modulation and Space Vector Pulse Width Modulation (SVPWM) for Power Conditioning in distribution systems. The Discrete PWM has the function of voltage stability, and harmonic suppression. The reference current can be calculated by 'd-q' transformation. In SVPWM technique, the Active Power Filter (APF) reference voltage vector is generated instead of the reference current, and the desired APF output voltage is generated by SVPWM. The THD will be decreased significantly by SVPWM technique than the Discrete PWM technique based Hybrid filters. Simulations are carried out for the two approaches by using MATLAB, it is observed that the %THD has been improved from 1.79 to 1.61 by the SVPWM technique.

KEYWORDS: Discrete PWM Technique, Hybrid Active Power Filter, Reference Voltage Vector, Space Vector Pulse Width Modulation (SVPWM), Total Harmonic Distortion (THD), Voltage Source Inverter (VSI).

I. INTRODUCTION

High power non-linear and time varying loads, such as rectifiers, office equipments like computers and printers, and also adjustable speed drives causes undesirable phenomena in the operation of power systems like harmonic pollution and reactive power demand [1-2]. The application of passive tuned filters creates new system resonances which are dependent on specific system conditions. In addition, passive filters often need to be significantly overrated to account for possible harmonic absorption from the power system. Passive filter ratings must be co-ordinate with reactive power requirements of the loads and it is often difficult to design the filters to avoid leading power factor operation for some load conditions [3-4]. Parallel active filters have been recognized as a viable solution to current harmonic and reactive power compensation. Various active power filter configurations and control strategies have been proposed and developed in the last decade in order to reduce these undesirable phenomena. Active filters have the advantage of being able to compensate for harmonics without fundamental frequency reactive power concerns. This means that the rating of the active power can be less than a comparable passive filter for the same non-linear load and the active filter will not introduce system resonances that can move a harmonic problem from one frequency to another. The active filter concept uses power electronics to produce harmonic current components that cancel the harmonic current components from the non-linear loads.

The active filter uses power electronic switching to generate harmonic currents that cancel the harmonic currents from a non-linear load. The active filter configuration investigated in this paper is

based on a discrete pulse-width modulation and pulse-width modulated (PWM) voltage source inverter based filters.

Among the various topologies the shunt active filter based on Voltage Source Inverter (VSI) is the most common one because of its efficiency [5]. The performance of active filter depends on the adoptive control approaches. There are two major parts of an active power filter controller. The first is that determines the reference current of APF and maintains a stable DC bus voltage. Various current detection methods, such as instantaneous reactive power theory, synchronous reference frame method, supplying current regulation, etc., are presented. The commonness of these methods is the request for generating reference current of Active Power Filter (APF), either with the load current or the mains current. The second is that controls the VSI to inject the compensating current into AC mains. The commonness of these methods is to control VSI with the difference between real current and reference current.

In discrete PWM technique based hybrid filters, the system has the function of voltage stability, and harmonic suppression. The reference current can be calculated by 'd-q' transformation [6-7]. In pulse-width modulated (PWM) voltage source inverter based filter differs from previously discussed approach in the following ways: a) To generate APF reference voltage vector instead of reference current; b) to generate desired APF output voltage by Space Vector Pulse Width Modulation (SVPWM) [8-9] based on generated reference voltage. Therefore, the proposed method is simple and easy to carry out. This paper discussed the basic principle of this method in detail and proved its validity by simulation results.

II. PROPOSED CONTROL METHODS

2.1. Using Discrete PWM Technique Based Hybrid Filter

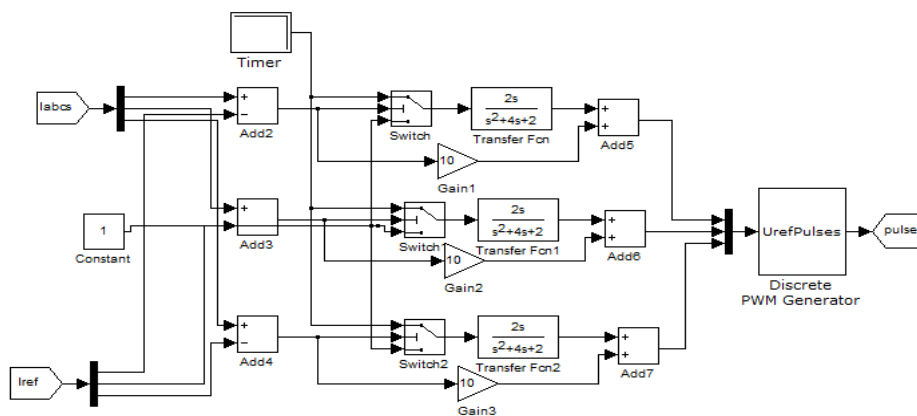


Figure 2.1. Simulation circuit of integral controller with Discrete PWM Generator

The Figure 2.1 is the integral controller used to generate the PWM pulses, which are generated based on the error produced by comparing the reference current and the source current. The differences calculated along with the gains are sent to discrete PWM generator and the resultant PWM pulses are given to the IGBT bridge for controlling.

2.2. d-q Transformation

The abc_to_dq0 Transformation block computes the direct axis, quadratic axis, and zero sequence quantities in a two-axis rotating reference frame for a three-phase sinusoidal signal. The following transformation is used

$$V_d = \frac{2}{3}(V_a \sin(\omega t) + V_b \sin(\omega t - 2\pi/3) + V_c \sin(\omega t + 2\pi/3)) \quad (2.1)$$

$$V_q = \frac{2}{3}(V_a \cos(\omega t) + V_b \cos(\omega t - 2\pi/3) + V_c \cos(\omega t + 2\pi/3)) \quad (2.2)$$

$$V_0 = \frac{1}{3}(V_a + V_b + V_c) \quad (2.3)$$

Where ω = rotation speed (rad/s) of the rotating frame.

The transformation is the same for the case of a three-phase current; which can be obtained by replacing the V_a , V_b , V_c , V_d , V_q , and V_0 variables with the I_a , I_b , I_c , I_d , I_q , and I_0 variables. This block

can be used in a control system to measure the positive-sequence component V_1 of a set of three-phase voltages or currents. The V_d and V_q (or I_d and I_q) then represent the rectangular coordinates of the positive-sequence component.

2.3. Control Method Using SVPWM Based Hybrid Filter

The main section of the APF shown in Figure 2.2 is a forced-commutated VSI connected to D.C capacitor [10]. Considering that the distortion of the voltage in public power network is usually very low, it can be assumed that the supply voltage is ideal sinusoidal and three-phase balanced as shown below:

$$\begin{aligned} V_{sa} &= V_s \sin(\omega t) \\ V_{sb} &= V_s \sin(\omega t - 2\pi/3) \\ V_{sc} &= V_s \sin(\omega t + 2\pi/3) \end{aligned} \quad (2.4)$$

Where V_s is the supply voltage amplitude.

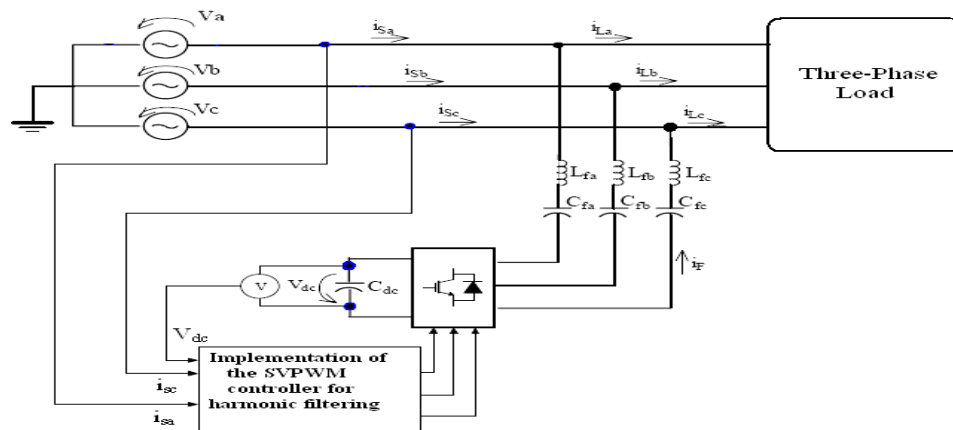


Figure 2.2. Configuration of an APF using SVPWM

It is known that the three-phase voltages $[v_{sa} \ v_{sb} \ v_{sc}]$ in a-b-c can be expressed as two-phase representation in d-q frame by Clark's transformation and it is given by

$$\overline{Vs} = \begin{bmatrix} Vd \\ Vq \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} Vsa \\ Vsb \\ Vsc \end{bmatrix} \quad (2.5)$$

It is possible to write equation (1.2) more compactly as

$$[\overline{Vs}] = \frac{2}{3} (V_{sa} a^0 + V_{sb} a^1 + V_{sc} a^2) = V_{sd} + jV_{sq} = V_s \angle \theta^s \quad (2.6)$$

Where $a = e^{j\frac{2\pi}{3}}$, so balanced three-phase set of voltages is represented in the stationary reference frame by a space vector of constant magnitude, equal to the amplitude of the voltages, and rotating with angular speed $\omega = 2\pi f$.

As shown in Figure 2.2, the shunt APF takes a three-phase voltage source inverter as the main circuit and uses capacitor as the energy storage element on the DC side to maintain the DC bus voltage V_{dc} constant. Figure 2.3 shows the per-phase (Phase A) equivalent circuit of the system.

2.4. Compensation Principle

In the Figure 2.3, $v_{fa,1}$ and $v_{fa,h}$ denote the output fundamental and harmonic voltages of the inverter, respectively. These voltage sources are connected to a supply source (v_{sa}) in parallel via a link inductor L_f and capacitor C_f . The supply current i_{sa} is forced to be free of harmonics by appropriate voltages from the APF and the harmonic current emitted from the load is then automatically compensated.

It is known from Figure 2.3, that only fundamental component is taken into account, the voltages of the AC supply and the APF exist the following relationship in the steady state

$$\overline{V_s} = L_f \cdot \frac{d\overline{I_{f1}}}{dt} + \frac{1}{C_f} \int \overline{I_{f1}} dt + \overline{V_{f1}} \quad (2.7)$$

Where \bar{V}_s is the supply voltage, \bar{I}_{f1} is the fundamental current of APF, \bar{V}_{f1} is the fundamental voltage of APF, and above variables are expressed in form of space vector.

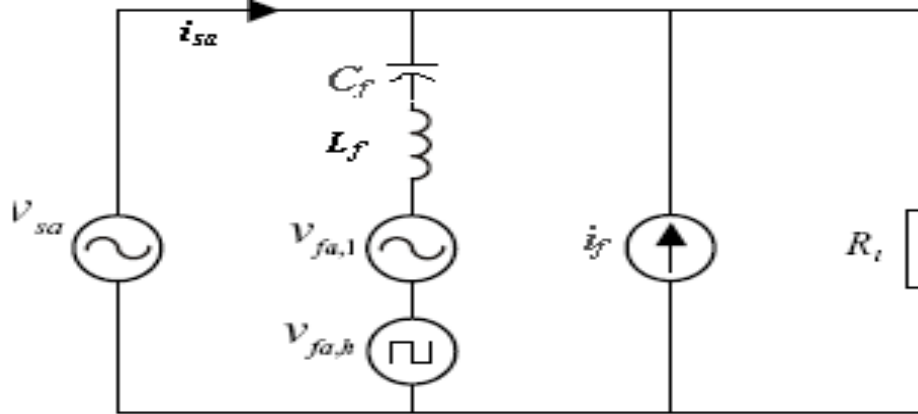


Figure 2.3. Equivalent circuit of a simple power system together with the APF

The APF is joined into the network through the inductor L_f and C_f . The function of these is to filter higher harmonics nearly switching frequency in the current and to link two AC voltage sources of the inverter and the network. So the required inductance and capacitance can just adopt a small value. Then the total reactance caused by inductor and capacitor for the frequency of 50Hz, and the fundamental voltages across the link inductors and capacitors are also very small, especially compared with the mains voltages. Thus the effect of the voltage of the link inductor and capacitor is neglected. So the following simplified voltage balanced equation can be obtained from equation (2.7).

$$\bar{V}_s = \bar{V}_{f1} \quad (2.8)$$

The control object of APF is to make the supply current sinusoidal and in phase with the supply voltage. Thus the nonlinear load and the active power filter equals to a pure resistance load R_s , and the supply voltage and the supply current satisfy the following equation:

$$\bar{V}_s = R_s \cdot I_s \quad (2.9)$$

Where $\bar{I}_s = \frac{2}{3}(i_{sa}a^0 + i_{sb}a^1 + i_{sc}a^2) = I_{sd} + jI_{sq} = I_s \angle \theta_i$

Then the relationship between I_s and the supply voltage amplitude V_s is

$$V_s = R_s \cdot I_s \quad (2.10)$$

Substituting (2.9), (2.10) into (2.8) results in

$$\bar{V}_{f1} = \frac{V_s}{I_s} \bar{I}_s \quad (2.11)$$

Equation (2.11) describes the relationship between the output fundamental voltage of APF, the supply voltage and the supply current, which ensure that the APF operate normally[11-12]. However, for making the APF normally achieving the required effect, the DC bus voltage V_{dc} has to be high enough and stable. In the steady state, the power supplied from the supply must be equal to the real power demanded by the load, and no real power passes through the power converter for a lossless APF system. Hence, the average voltage of DC capacitor can be maintained at a constant value. If a power imbalance, such as the transient caused by load change, occurs, the DC capacitor must supply the power difference between the supply and the load, the average voltage of the DC capacitor is reduced. At this moment, the magnitude of the supply current must be enlarged to increase the real power delivered by the supply. On the contrary, the average voltage of the DC capacitor rises, and the supply current must be decreased. Therefore, the average voltage of the DC capacitor can reflect the real power flow information. In order to maintain the DC bus voltage as constant, the detected DC bus voltage is compared with a setting voltage. The compared results is fed to a PI controller, and amplitude control of the supply current I_s can be obtained by output of PI controller.

The Figure 2.4 shows the block diagram of active filter controller implemented for reducing the harmonics with hybrid active filter system. In each switching cycle, the controller samples the supply currents i_{sa} , i_{sc} and the supply current i_{sc} is calculated with the equation of $-(i_{sa} + i_{sc})$, as the summation of three supply current is equal to zero. These three-phase supply currents are measured and transformed into synchronous reference frame (d-q axis) [13-14]. The fundamental component of

the supply current is transformed into DC quantities in the (d-q) axis and the supply current amplitude I_s generated by the PI controller with V_{dc} and V_{ref} , the reference value of the DC bus voltage. The obtained d-q axis components generate voltage command signal. By using Fourier magnitude block, voltage magnitude and angle is calculated from the obtained signal. These values are fed to the developed code and compared with the repeating sequence. Then the time durations T_1 , T_2 and T_0 , the on-time of V_1 , V_2 and V_0 are calculated. The generated switching actions are applied to the APF and power balancing of the filter takes place.

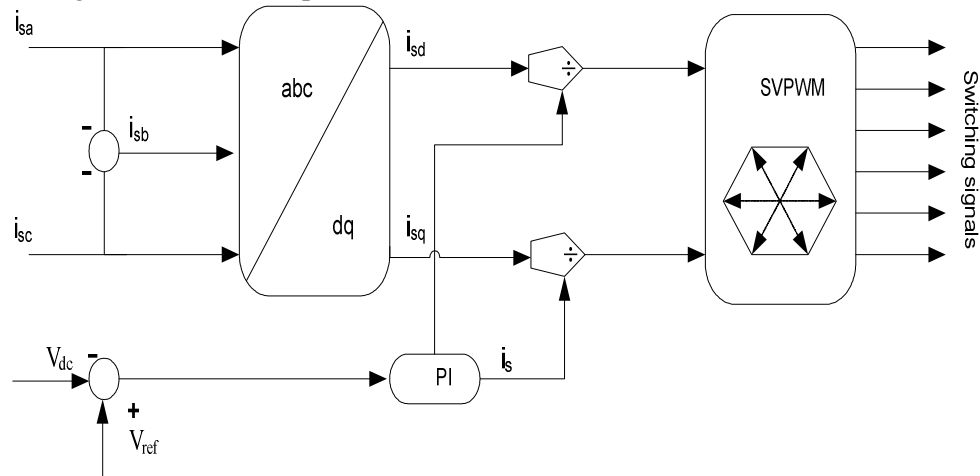


Figure 2.4. Control block diagram of proposed algorithm

III. RESULTS AND DISCUSSIONS

3.1. Discrete PWM Technique Based Hybrid Shunt Active Power Filter

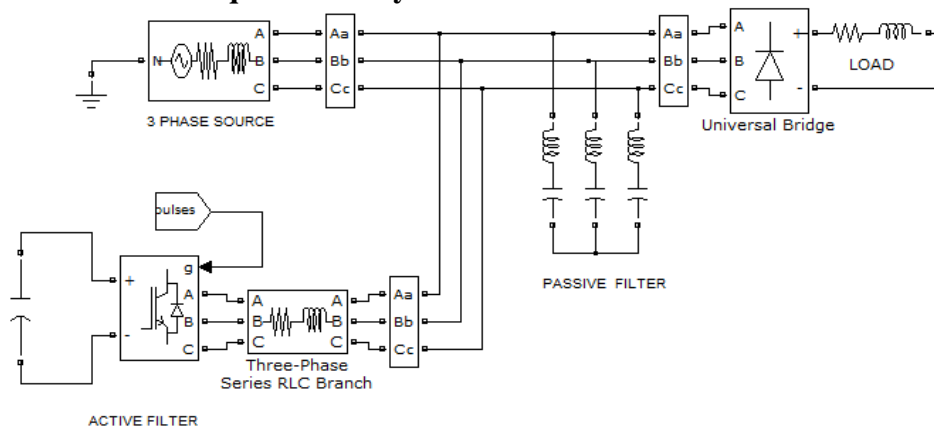


Figure 3.1. Simulation circuit of hybrid shunt active power filter

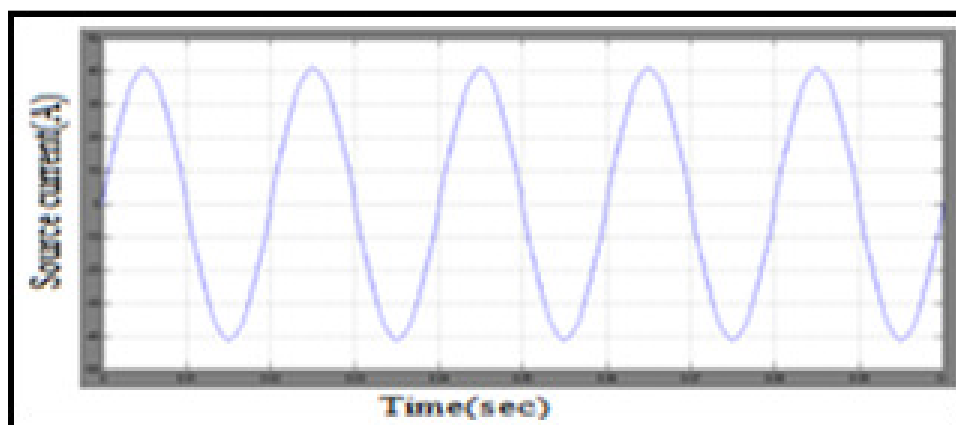


Figure 3.2. Source current waveform with hybrid filter

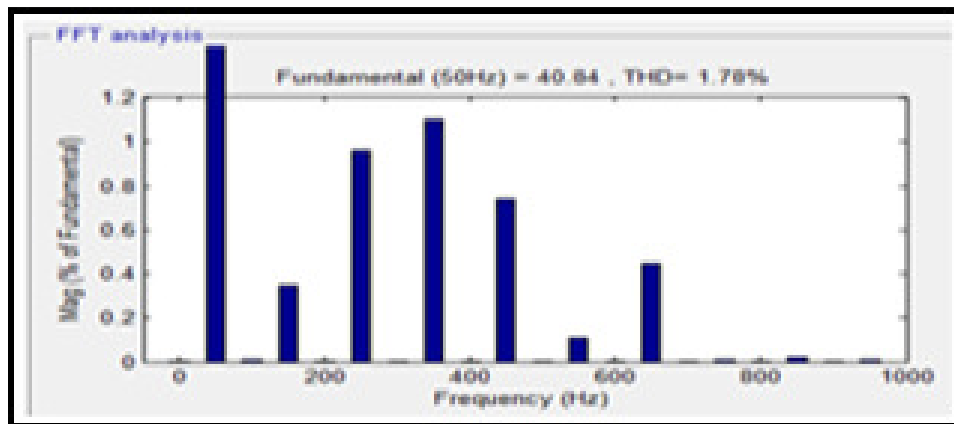


Figure 3.3. FFT analysis of source current with hybrid filter

The circuit shown in Figure 3.1 consists of both active and passive filters, the main purpose of this hybrid filter is that it reduces the harmonic content to a larger extent compared with the above methods. The lower and higher order harmonics are reduced by the passive filter and the other order harmonics are reduced by the active filter. In the case of hybrid filters, the wave form appears in a sinusoidal shape and the distortions are less as shown in the Figure 3.2 compared with the previous techniques. So, the hybrid filter is preferred due to its better performance characteristics. From the FFT analysis of source current with hybrid filter, It is observed that the %THD is 1.78.

3.2. SVPWM Technique Based Hybrid Shunt Active Power Filter

The developed control method for three-phase shunt APF is simulated in MATLAB/ Simulink. Firstly, the three-phase supply currents are sensed and transformed into synchronous reference frame (d-q) axis. The fundamental component of the supply current is transformed into DC quantities in the (d-q) axis and the supply current amplitude I_s generated by the PI controller. The obtained d-q axis components generate voltage command signal. By using Fourier magnitude block, voltage magnitude and angle is calculated from the obtained signal. These values are fed to the developed code and generated switching actions are applied to the APF. Thus, power balancing of the filter takes place. Further, the performance with different type of loads is presented. The complete simulation model of APF with different type of loads is shown in Figure 3.5 and Figure 3.8. For an input supply voltage of 230V (rms) and switching frequency of 5kHz, the simulation results before and after power balancing are shown.

3.2.1. For Balanced Linear Load

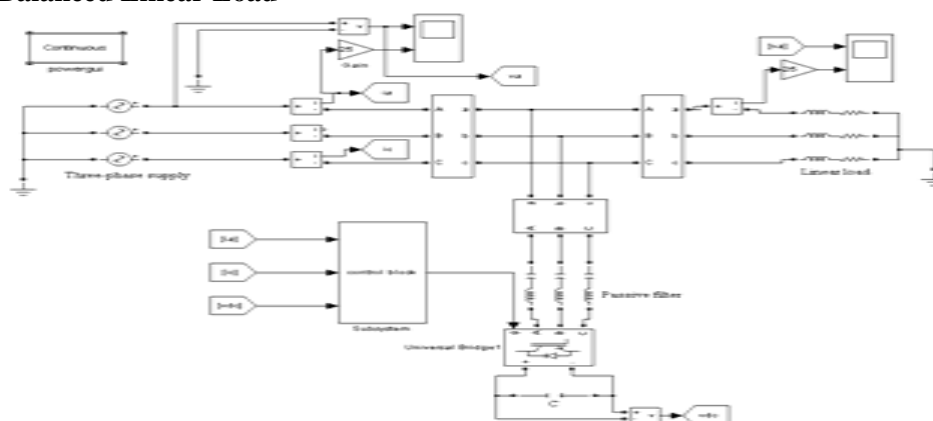


Figure 3.5. Simulation model of APF with linear load

The Figure 3.6 shows the simulation results of the APF when load is three-phase balanced RL load. Figure 3.6 (a) is the waveforms of the phase-A supply voltage and the load current before compensation. Figure 3.6 (b) is the waveforms of the phase-A supply voltage and the supply current after compensation.

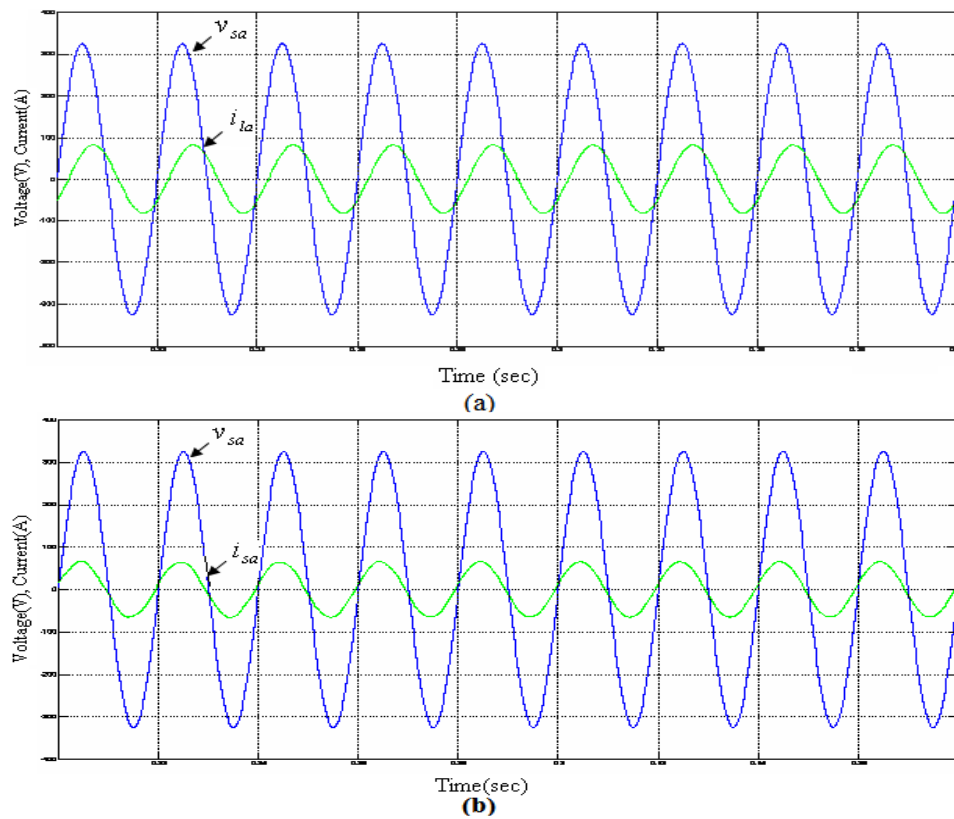


Figure 3.6. Simulation results of balanced linear load

(a) The phase-A supply voltage and load current waveforms

(b) The phase-A supply voltage and supply current waveforms

3.2.2. For Unbalanced Linear Load

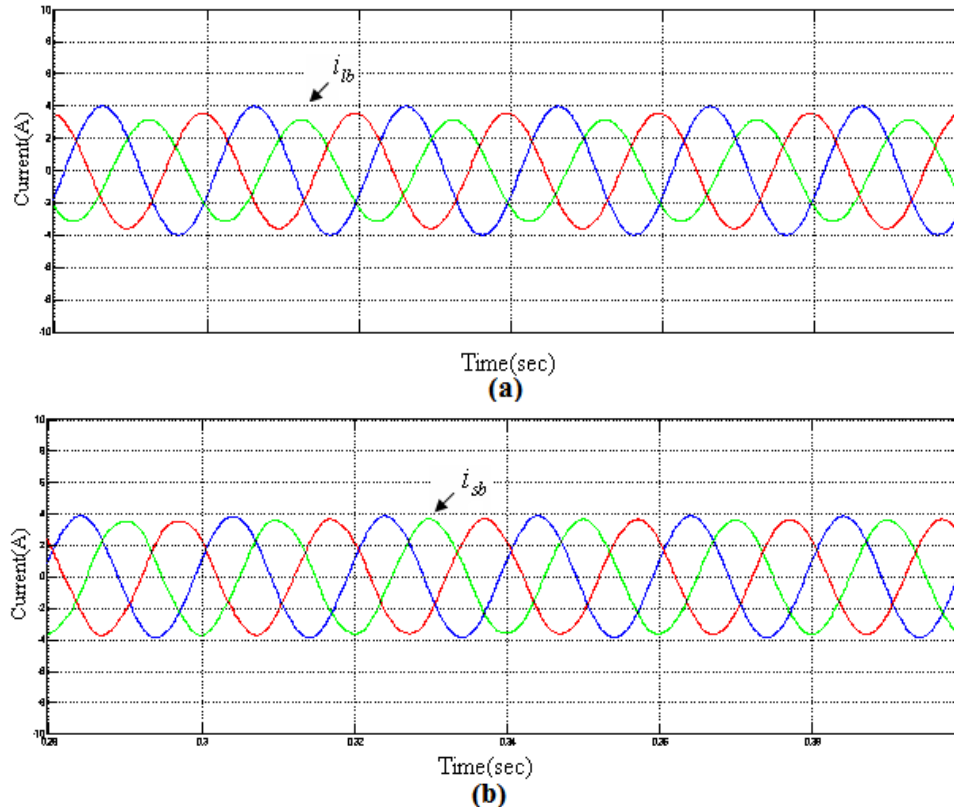


Figure 3.7. Simulation results of unbalanced linear load

(a) Three-phase load current waveforms

(b) Three-phase supply current waveforms

The Figure 3.7 shows the simulation results of APF when three-phase unbalanced RL load is considered. Figure 3.7 (a) is the waveforms of the three-phase load current before compensation. Figure 3.7 (b) is the waveforms of the three-phase mains current after compensation. From the figures, it can be seen that APF controller can remedy the system unbalance.

3.2.3. For Non-Linear Load with Resistance

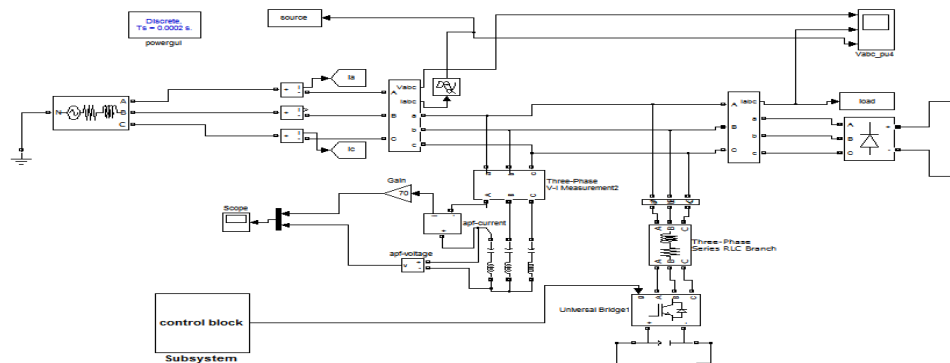


Figure 3.8. Simulation model of APF with non-linear load

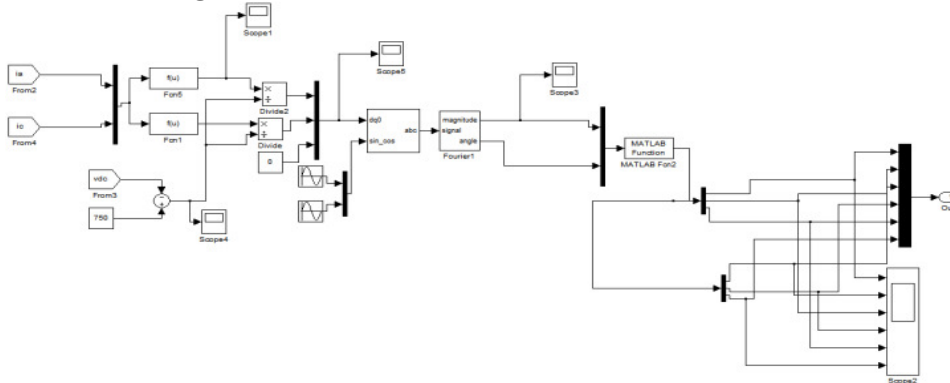
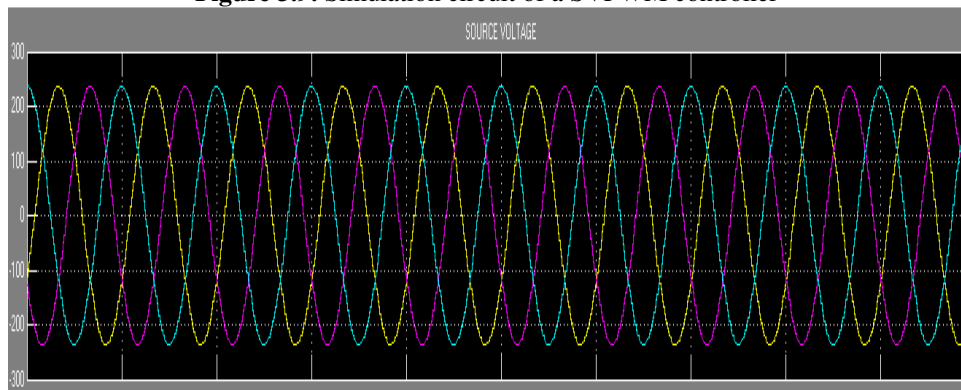
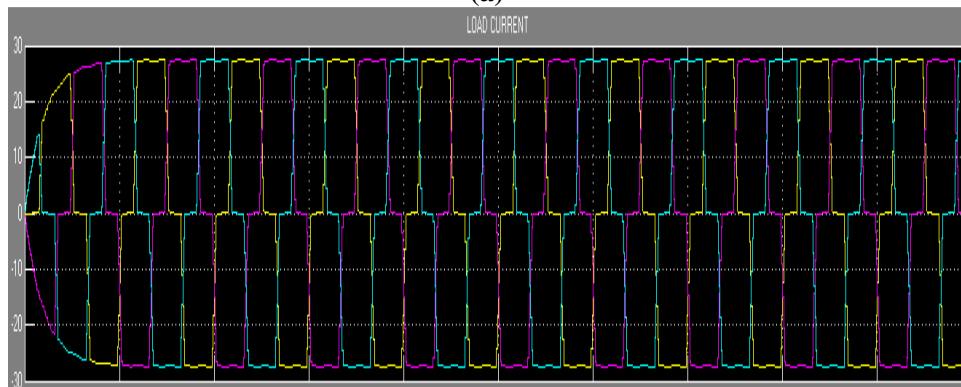


Figure 3.9. Simulation circuit of a SVPWM controller



(a)



(b)

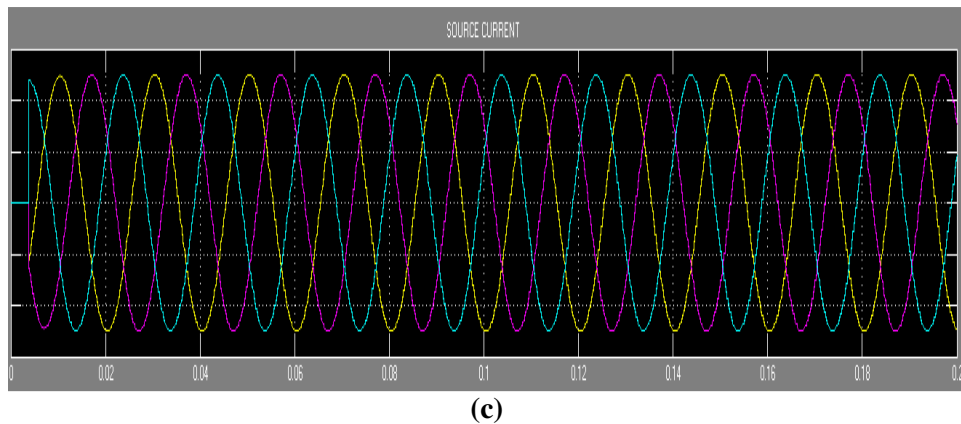


Figure 3.10. Simulation results of non-linear load
 (a) The three-phase source voltage waveforms
 (b) The three-phase load current waveforms
 (c) The three-phase source current waveforms

The Figure 3.8 and Figure 3.9 shows the simulation model of APF with non-linear load and its SVPWM controller respectively. The Figure 3.10 shows the behaviour of the APF when the non-linear load is a three-phase diode bridge rectifier with resistance load. Figure 3.10 (a) is the waveforms of the source phase voltage. Figure 3.10 (b) is the wave forms of the load current before compensation. Figure 3.10 (c) is the waveforms of the supply current after compensation.

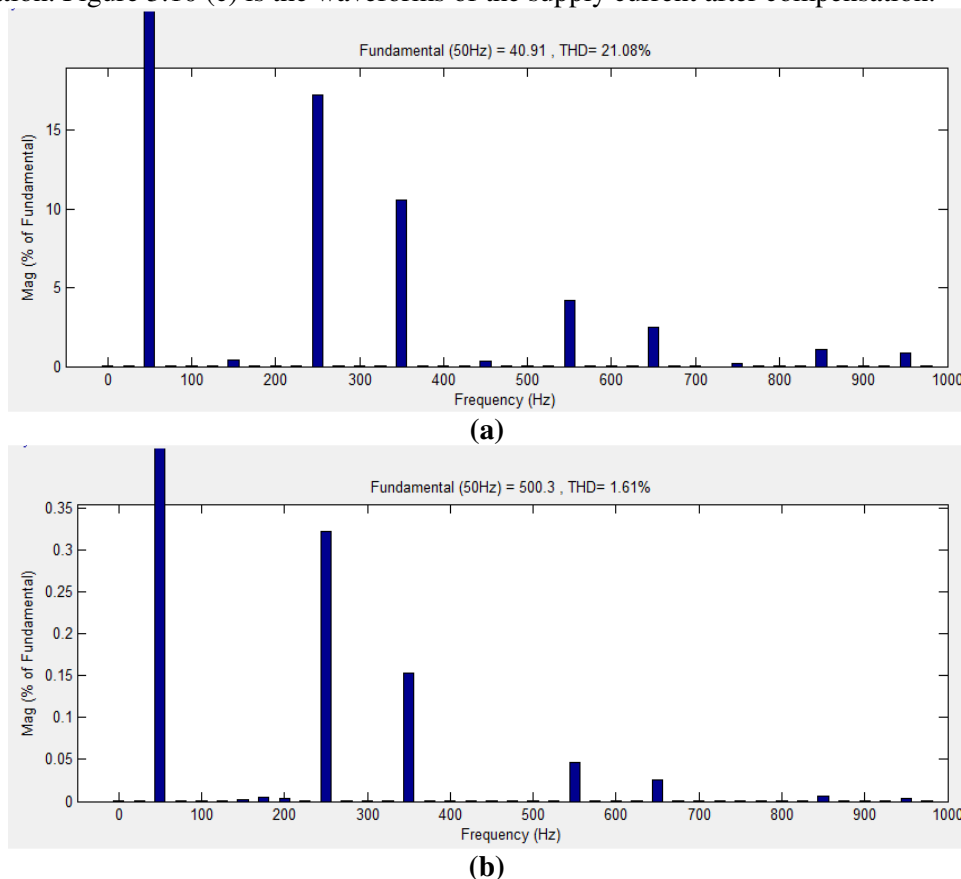


Figure 3.11. Harmonic spectrum of non-linear load
 (a) The phase-A load current harmonic spectrum
 (b) The phase-A source current harmonic spectrum

The Figure 3.11 shows the simulation of harmonic spectrum of APF when the non-linear is a three-phase diode bridge rectifier with resistance load. Figure 3.11 (a) is the harmonic spectrum of the current before compensation on the load side. Figure 3.11 (b) is the harmonic spectrum of the current after compensation on the source side. The harmonic spectrum of the load current shows that magnitude of the 5th, 7th, 11th and 13th harmonics is very large. The harmonic spectrum of the source

current shows that magnitude of the 5th, 7th, 11th and 13th harmonics are evidently reduced after compensation. The load current Total Harmonic Distortion (THD) is 21.08%, while the supply current THD is 1.61%. It should be noted that the higher frequency harmonics caused by APF in mains current can be canceled easily by a small passive filter, and there are pulses in main current at the points, where di/dt of load current is large, because fixed switching frequency restrict the tracking capability of APF.

IV. CONCLUSIONS

In this paper, a control methodology for the APF using Discrete PWM and SVPWM is proposed. These methods require a few sensors, simple in algorithm and are able to compensate harmonics and unbalanced loads. The performance of APF with these methods is done in MATLAB/Simulink. The algorithm will be able to reduce the complexity of the control circuitry. The harmonic spectrum under non-linear load conditions shows that reduction of harmonics is better. Under unbalanced linear load, the magnitude of three-phase source currents are made equal and also with balanced linear load the voltage and current are made in phase with each other. The simulation study of two level inverter is carried out using SVPWM because of its better utilization of DC bus voltage more efficiently and generates less harmonic distortion in three-phase voltage source inverter. This SVPWM control methodology can be used with series APF to compensate power quality distortions.

From the simulated results of the filtering techniques, it is observed that Total Harmonic Distortion is reduced to an extent by the SVPWM Hybrid filter when compared to the Discrete PWM filtering technique i.e. from 1.78% to 1.61%.

REFERENCES

- [1] EI-Habrouk. M, Darwish. M. K, Mehta. P, "Active Power Filters-A Rreview," Proc.IEE-Elec. Power Applicat., Vol. 147, no. 5, Sept. 2000, pp. 403-413.
- [2] Akagi, H., "New Trends in Active Filters for Power Conditioning," IEEE Trans. on Industry applications, Vol. 32, No. 6, Nov-Dec, 1996, pp. 1312-1322.
- [3] Singh.B, Al-Haddad.K, Chandra.A, "Review of Active Filters for Power Quality Improvement," IEEE Trans. Ind. Electron., Vol. 46, No. 5, Oct, 1999, pp. 960-971.
- [4] Ozdemir.E, Murat Kale, Sule Ozdemir, "Active Power Filters for Power Compensation Under Non-Ideal Mains Voltages," IEEE Trans. on Industry applications, Vol.12, 20-24 Aug, 2003, pp.112-118.
- [5] Dan.S.G, Benjamin.D.D, Magureanu.R, Asimionoaei.L, Teodorescu.R, Blaabjerg.F, "Control Strategies of Active Filters in the Context of Power Conditioning," IEEE Trans. on Ind. applications, Vol.25,11-14 Sept-2005, pp.10-20
- [6] Mr H. Bouchikha, M. Ghers, "Three Phase Shunt Hybrid Filters for the Current Harmonics Suppression and the Reactive Power Compensation", European Journal of Scientific Research, ISSN 1450-216X, Vol.24 No.4 (2008), pp.580-590.
- [7] João Afonso, Maurício Aredes, Edson Watanabe, Júlio Martins, "Shunt Active Filter for Power Quality Improvement", International Conference UIE 2000, Portugal, 1-4 November 2000, pp. 683-691.
- [8] Atif Iqbal, Lamine.A, Imtiaz.Ashraf, Mohibullah, "MATLAB Model of Space Vector PWM for Three-Phase Voltage Source Inverter," universities power engineering conference, 2006, UPEC'06, proceedings of the 41st international Vol. 3, 6-8 Sept. 2006, pp:1096-1100.
- [9] Rathnakumar.D, LakshmanaPerumal, Srinivasan.T, "A New Software Implementation of Space Vector PWM," IEEE Trans. Power Electron., Vol.14,8-10 April 2005, pp.131-136.
- [10] Wang Jianze, Peng Fenghua, Wu Quitao, Ji Yanchao, "A novel control method for shunt active power filters using svpwm," IEEE Trans. on Industry applications, vol.1, 3-7 Oct, 2004, pp.134-139.
- [11] Roger C. Dugon, Mark F. McGranaghan, H.Wayne Beaty, Surya Sontoso, "Power Systems Quality", TATA Mc Graw-Hill, 2nd Edition.
- [12] Mr. Sivanagaraju, "Electrial Distribution Systems", a text book on power distribution.
- [13] G. Bhuvaneswari and M.G.Nair, " Three-Phase Hybrid Shunt Filters for Power Quality Improvement," Journal of Power Electronics, Vol. 7, No. 3, pp.257-264, 2007.
- [14] Mr H. Bouchikha, M. Ghers, "Three Phase Shunt Hybrid Filters for the Current Harmonics Suppression and the Reactive Power Compensation", European Journal of Scientific Research, ISSN 1450-216X Vol.24 No.4 (2008), pp.580-590.

Authors Biographies

Jarupula Somlal, at present is working as an Assistant Professor in the department of EEE, K.L.University, Guntur, Andhra Pradesh, India. He received B.Tech, degree in Electrical and Electronics Engineering from J.N.T.University, Hyderabad, A.P, India, M.Tech.,(Electrical Power Engineering) from J.N.T.University, Hyderabad, A.P, India and currently working towards the Doctoral degree in Electrical & Electronics Engineering at Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. He published 3 papers in international Conferences and presented various papers in national and International conferences. His current research interests include active filtering for power conditioning, Fuzzy Logic and ANN applications to power quality.



Venu Gopala Rao Mannam, FIE, Member IEEE, at present is Professor & Head, department of Electrical & Electronics Engineering, K L University, Guntur, Andhra Pradesh, India. He received B.E. degree in Electrical and Electronics Engineering from Gulbarga University in 1996, M.E (Electrical Power Engineering) from M S University, Baroda, India in 1999, M.Tech (Computer Science) from JNT University, India in 2004 and Doctoral Degree in Electrical & Electronics Engineering from J.N.T.University, Hyderabad, India in 2009. He published more than 20 papers in various National, International Conferences and Journals. His research interests accumulate in the area of Power Quality, Distribution System, High Voltage Engineering and Electrical Machines.



COMPARISONS AND LIMITATIONS OF BIOHYDROGEN PRODUCTION PROCESSES: A REVIEW

Karthic Pandu and Shiny Joseph

Department of Chemical Engineering, National Institute of Technology Calicut,
Kozhikode, Kerala State, India,

ABSTRACT

Hydrogen gas can be produced by conventional methods such as thermo-chemical gasification, pyrolysis, solar gasification and supercritical conversion. Hydrogen productions through biological methods are an attractive and alternate method for the replacement of fossil fuel in future. In this review, the major biological processes discussed for hydrogen production are bio-photolysis of water by algae, dark fermentation, photo-fermentation of organic materials and the sequential dark and photo-fermentation processes. Major constraints in dark and photo-fermentative hydrogen production include the raw material cost, lower hydrogen yield and rate of hydrogen production. To overcome those constraints, intensive research works are strongly recommended to be carried out on the advancement of these processes. The review showed effective utilization of low cost substrate such as agricultural, food industry wastes and effluents such as dairy industry wastewater for hydrogen production with inexpensive energy generation and simultaneous wastewater treatment. It reveals that the hydrogen yield could be even achieved greater with the effective pretreatment methods of inoculum and substrates. This review explores the recent status and developments that have been made to improve the hydrogen production particularly with pretreatment methods and gave an outline about the unit cost of various hydrogen production processes.

KEYWORDS: Biohydrogen, Bio-Photolysis, Dark-Fermentation, Photo-Fermentation and Pretreatment Methods.

I. INTRODUCTION

Hydrogen becomes a significant and alternate energy carrier to fossil fuels because of its property of clean, renewable, high energy content and does not contribute to environmental problems such as greenhouse effect that leads to global warming. Recent reviews on hydrogen indicated that the worldwide need on hydrogen is increasing with a growth rate of nearly 12% per year for the time being and contribution of hydrogen to total energy market will be 8-10% by 2025 [1]. Conventional hydrogen gas production methods are energy intensive processes requiring high temperatures (>840°C). Electrolysis of water may be the cleanest technology for hydrogen gas production. However, electrolysis should be used in areas where electricity is inexpensive since electricity costs account for 80% of the operating cost of H₂ production. At present the total annual world hydrogen production is around 368 trillion cubic meters (2008). Of this amount, about 40% is used in the chemical industry, 40% in refineries and the other 20% in a large variety of processes [2] including its use as energy carrier. In 2005, 48% of the global demand for hydrogen was produced from steam reforming of natural gas, about 30 % from oil /naphtha reforming from refinery/chemical industrial off- gases, 18% from coal gasification, 3.9% from water electrolysis and 0.1% from other sources [3]. Due to increasing trend of hydrogen demand, development of cost effective and efficient hydrogen production technologies has gained significant attention. Biohydrogen production technology can

utilize the renewable energy sources like biomass for the generation of hydrogen, a cleanest energy carrier for the use of mankind. The hydrogen is considered as a dream fuel by virtue of the fact that it has high energy content per unit mass of any known fuel (142MJ/kg), is easily converted to electricity by fuel cells and on combustion it gives water as the only byproduct. Presently, 40 % H₂ is produced from natural gas, 30 % from heavy oils and naphtha, 18 % from coal, and 4 % from electrolysis and about 1 % is produced from biomass.

The advantages and disadvantages of various hydrogen production processes are outlined in Table 1[4]. Among various hydrogen production processes, biological method is known to be less energy intensive, for it can be carried out at ambient temperature and pressure [5]. Biological method mainly includes photosynthetic hydrogen production and fermentative hydrogen production.

Table 1: Comparison of various hydrogen production processes

Process	Advantages	Disadvantages
Solar gasification	Good hydrogen yield	Effective solar collector plates are required
Thermo-chemical gasification	Higher conversion can be achieved	Gas conditioning and tar removal is to be done
Pyrolysis	Gives carbonaceous material with oil, chemicals and minerals	Catalyst deactivation will occur
Supercritical conversion	Sewage sludge can be used easily, difficult by gasification	Selection of supercritical medium
Direct bio-photolysis	H ₂ can be produced directly from water and sunlight	Requires high intensity of light, low photochemical efficiency and O ₂ is inhibitory.
Indirect bio-photolysis	Blue green algae can produce hydrogen from water. It has the ability to fix N ₂ from atmosphere	Uptake hydrogenates are to be removed.
Photo-fermentation	A wide spectral energy can be used by photosynthetic bacteria.	O ₂ is inhibitory on nitrogenase enzyme and light conversion efficiency is low.
Dark fermentation	It can produce H ₂ without light. No oxygen limitations and can produce several metabolites as by-products. Various substrates can be used in this anaerobic process.	Relatively lower H ₂ yield. At higher H ₂ yield, process becomes thermodynamically unfavorable.
Two-stage fermentation	Can produce relatively higher H ₂ yield. By-products (metabolites) can be efficiently converted to H ₂ .	Requires continuous light source which is difficult for large scale processes.

Even though photosynthetic hydrogen production is a theoretically perfect process with transforming solar energy into hydrogen by photosynthetic bacteria, applying it to practice is difficult due to the low utilization efficiency of light and difficulties in designing the reactors for hydrogen production [6]. However, fermentative hydrogen production has the advantages of rapid hydrogen production rate and simple operation. Moreover, it can use various organic wastes as substrate for fermentative hydrogen production. Thus, compared with the photosynthetic hydrogen production, fermentative hydrogen production is more feasible and thus widely used. In addition, it is of great significance to produce hydrogen from organic wastes by fermentative hydrogen production, because it can not only treat organic wastes, but also produce very clean energy. Therefore fermentative hydrogen production has been received increasing attention in recent years. For cost effective, we need to go for the utilization of agricultural and some food industry effluents such as cheese whey, olive mill and baker's yeast industry wastewaters for hydrogen production. It also provides inexpensive energy generation with simultaneous wastewater treatment.

II. BIOHYDROGEN PRODUCTION METHODS

2.1. Bio-photolysis

2.1.1. Direct process

The action of light on a biological system that results in the dissociation of a substrate, usually water, to produce hydrogen is referred to as biophotolysis. A direct biophotolysis of H_2 production is a biological process which utilizes solar energy and photosynthetic systems of algae to convert water into chemical energy.

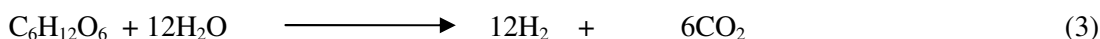
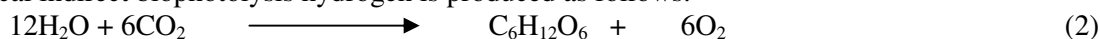
The concept of the biophotolysis of water with the formation of oxygen and hydrogen is the bringing together of two biological fields of scientific endeavor, each of which has made phenomenal progress during the last two decades. The two areas of progress referred to are: (1) a greater understanding of the molecular events which occur in photosynthesis, and (2) a greater understanding of molecular events in microbial metabolism, although thermodynamically feasible, heretofore not much thought has been given to the possibility of bio-photolysis. The photosynthetic system which consists of two photo-systems operating in series can, by capturing two quanta of radiant energy, place an electron from the water-oxygen couple (0.8 volts pH 7.0) to a negative value as much as -0.7 volt which is 0.3 volts more negative than the hydrogen electrode. A minimum of eight quanta of radiant energy are required for the following photosynthetic equation:



where A is an electron acceptor. A variety of compounds may serve as Hill reagents. For the purpose of employing these photosynthetic electrons for the reduction of protons to hydrogen by the action of a bacterial hydrogenase, the acceptor must have an oxidation-reduction potential near the potential of the hydrogen electrode and in its reduced state serve as a substrate for the hydrogenase. In this reaction oxygen produced by the photosynthesis strongly inhibits the hydrogen production [7]. Inhibition is not due to the oxygen inactivation of hydrogenase, mainly due to the reaction of oxygen with the photo system (ferredoxin or hydrogenase). Development by biotechnology of an oxygen stable Hydrogenase reaction is not plausible on thermodynamic and other backgrounds. To overcome oxygen inhibition, photosynthesis process needs to produce oxygen absorbers (eg; glucose-oxidase). Biological hydrogen production is often conducted in two stages under different atmospheric conditions, the first stage for cell growth followed by the second stage for hydrogen evolution. Nitrogen starvation is often at the end of the growth stage as an efficient metabolic stress to induce the nitrogenase activity. It was clear that a necessary technology breakthrough need to be attained for hydrogen productivity by nitrogen fixing cyanobacteria. However, direct biophotolysis, though limited by its relatively low hydrogen productivity, provides a working model for hydrogen production from water and sunlight energy. Therefore it is necessary to develop knowledge and technical innovations in hydrogen enzymes, electron carriers that would lead to a bio-mimetic water photolysis system avoids intrinsic incompatibility of simultaneous hydrogen and oxygen evolution and splits water into separated gas streams.

2.1.2. Bio-photolysis-Indirect process

The most credible processes for future applied research and development are those which couple separate stages of microalgal photosynthesis and fermentations ('indirect biophotolysis'). These involve fixation of CO_2 into storage carbohydrates (e.g. starch in green algae, glycogen in cyanobacteria) followed by their conversion to H_2 by the reversible hydrogenase, both in dark and possibly light-driven anaerobic metabolic processes [8]. In indirect biophotolysis, the problem of sensitivity of the H_2 evolving process to O_2 is usually circumvented by separating O_2 and H_2 [9]. In a typical indirect biophotolysis hydrogen is produced as follows:



Based on a preliminary engineering and economic analysis, biophotolysis processes must achieve close to an overall 10% solar energy conversion efficiency to be competitive with alternatives sources of renewable H_2 , such as photovoltaic electrolysis processes. Such high solar conversion efficiencies in photosynthetic CO_2 fixation could be reached by genetically reducing the number of light harvesting chlorophylls and other pigments in microalgae. Similarly, greatly increased yields of H_2 from dark fermentation by microalgae could be obtained through application of the techniques of metabolic engineering. Another challenge is to scale-up biohydrogen processes with economically viable bioreactors. Solar energy driven microalgae processes for biohydrogen production are

potentially large-scale, but also involve long-term and economically high-risk. In the nearer-term, it may be possible to combine microalgal H_2 production with wastewater treatment.

2.2. Dark Fermentation

Dark fermentation is the fermentative conversion of organic substrate to biohydrogen. It is a complex process manifested by diverse group of bacteria by a series of biochemical reactions. Fermentative/hydrolytic microorganisms hydrolyze complex organic polymers to monomers which further converted to a mixture of lower molecular weight organic acids and alcohols by necessary H_2 producing acidogenic bacteria. Utilization of wastewater as a potential substrate for biohydrogen production has been drawing considerable interest in recent years especially in dark fermentation process. Industrial wastewater as fermentative substrate for H_2 production addresses most of the criteria required for substrate selection viz., availability, cost and biodegradability [10]. Chemical wastewater [11], cattle wastewater [12] and starch hydrolysate wastewater [13] have been reported to produce biohydrogen apart from wastewater treatment from dark fermentation process using selectively enriched mixed culture under acidophilic conditions. Various wastewaters viz., paper mill wastewater [14], food processing wastewater [15], rice winery wastewater [16], distillery and molasses based wastewater [17]; wheat straw wastes [18] and palm oil mill wastewater were also studied as fermentable substrates for H_2 production along with wastewater treatment. Utilizing mixed culture is extremely important and well-suited to the non-sterile, ever-changing, complex environment of wastewater treatment. Some anaerobic mixed cultures cannot produce H_2 as it is rapidly consumed by the methane-producing bacteria. Successful biological H_2 production requires inhibition of H_2 consuming microorganisms, such as methanogens and pre-treatment of parent culture is one of the strategies used for selecting the requisite microflora. The physiological differences between H_2 producing bacteria and H_2 consuming bacteria (methanogenic bacteria) form the fundamental basis behind the development of various methods used for the preparation of H_2 producing seeds. When the inoculum was exposed to extreme environments such as high temperature, extreme acidity and alkalinity, spore forming H_2 producing bacteria such as *Clostridium* survived, but methanogens had no such capability. Also, hydrogen production from different agricultural wastes has been reviewed by Guo et al [19] given the present state of knowledge; further experimentation is required to better understand the impact on biohydrogen production performances of the compositions and the characteristics of different organic substrates. Pretreatment processes of agricultural waste also require specific investigation since the origins and compositions of the agricultural wastes determine which specific pretreatment is the most suitable.

In the Table 2, it was shown that higher hydrogen yield was achieved by utilizing mixed consortia rather than pure culture. Also mutant strain or genetically engineered microorganism will yield higher than the others. Though possessed significant advantages, the main challenge observed with fermentative H_2 production process was relatively low energy conversion efficiency from the organic source. Typical H_2 yields range from 1 to 2 mol of H_2 /mol of glucose, which resulted in 80-90% of the initial COD remaining in the wastewater in the form of various volatile organic acids (VFAs) and solvents [38]. Even under optimal conditions about 60-70% of the original organic matter remains in solution [39]. Usage of unutilized carbon sources present in acidogenic process for additional biogas production sustains the practical applicability of the process. One way to utilize the remaining organic matter in a usable form is to produce additional H_2 by terminal integration of photo-fermentative process for H_2 production and methane by integrating acidogenic process to terminal methanogenic process [40]. Baghchehsaree et al [21] reported the H_2 yield of 2.18mol/mol glucose by using *Thermotoga neapolitana*, whereas Cakir et al [25], reported 2.4mol/mol glucose from wheat straw wastes using anaerobic sludge. Though the earlier have used pure culture and substrate (glucose), it yields lesser than the later using anaerobic sludge. The possible reason is that the waste must be subjected to appropriate pretreatment before deploying it for fermentative hydrogen production to enrich the hydrogen producing bacteria, repressed the methanogenic activity. This enhances the higher yield, in dark fermentation, thereby paves the way for simultaneous waste treatment and inexpensive energy generation from the low cost substrate.

Perera et al [41] evaluated the net energy gain for dark fermentation processes using various substrates. Evaluation reported that the improvement in hydrogen yield at higher temperature is not

justified as the net energy gain not only declined with increase of temperature, but also was mostly negative for above the ambient temperature.

Table 2: Comparison of various processes on biohydrogen production by dark fermentation

System	Culture	Substrate	Volumetric H ₂ production rate, (mL ⁻¹ h ⁻¹)	H ₂ Yield, (molH ₂ /mol glucose ⁻¹)	Reference
Batch	<i>Thermotoga neapolitana</i>	Glucose	51	3.85	[20]
Batch	Anaerobic digester sludge	Glucose	-	2.18	[21]
Continuous	Mixed culture	Glucose	5100	2.1	[22]
Batch	Rice Rhizosphere microflora	Apple pomace wastes	-	2.3	[23]
Continuous	<i>C. acetobutylicum</i> ATCC 824	Glucose	1270	-	[24]
Batch	Heat treated anaerobic sludge	Acid-Hydrolyzed ground wheat	7.42	2.4	[25]
Batch	Defined consortium	Sewage sludge	-	35.54 mL g sludge ⁻¹	[26]
Batch	<i>Bacillus coagulans</i>	Glucose	-	2.28	[27]
Batch	<i>Klebsiella oxytoca</i> HP1	Glucose	350	3.6	[28]
Batch	Cow dung compost	Wheat straw wastes	10.14 mL g TVS ⁻¹ h ⁻¹	68.114 mL g TVS ⁻¹	[18]
Batch	<i>Clostridia</i> sp	Glucose	1496	1.7	[29]
Batch	<i>E. cloacae</i> IIT-BT 08	Sucrose	660	6.0	[30]
Batch	<i>Thermoanaerobacterium</i>	Starch in wastewater	1.9 ml/h	92	[31]
Batch	<i>C. pasteurianum</i>	Starch	4.2 mlh ⁻¹	106 ^b	[32]
Batch	Mixed culture	Glucose	-	2.1	[33]
Continuous	Anaerobic mixed consortia	Diary wastewater	-	0.122 ^a	[34]
Batch	Mixed culture	Sewage bio-solids	-	700 ^a	[35]
Continuous	Sludge from WTP	Rice winery wastewater	-	2.14	[16]
Batch	Sludge from WTP	Potato processing wastewater	-	6.0 ^a	[16]
Batch	<i>Thermoanaerobacterium</i>	Starch wastewater	-	92 ^b	[31]
Continuous	Activated sludge and digested sludge	Glucose	2360	1.16	[36]
Continuous	<i>T. kodakaraensis</i> KOD1	Starch	146 ^c	-	[37]

a - mmol H₂ gCOD⁻¹
b - mL H₂ g⁻¹ starch
c - mmol H₂ gdw⁻¹ h⁻¹

It was concluded to operate the dark fermentation at the near-ambient temperatures to maximize the net energy gain [41]. And this should be done by direct electricity production from the dark fermentation via microbial fuel cells. But this was contradictory to the earlier study by Wu et al [42], has achieved positive net energy gain of 0.8kJ/g COD in spite of high fermentation temperatures of

40° C. Reason for this higher achievement of net energy gain is that they utilized high sucrose concentration of 30g/L and high hydrogen producing bacteria [42].

2.2.1. Effect of Pretreatment Methods

Pre-treatment helps to accelerate the hydrolysis step, thus, reducing the impact of rate limiting step and augment the anaerobic digestion to enhance the H₂ generation [43]. Several pre-treatment procedures viz., heat-shock, chemical, acid, alkaline, oxygen-shock, load-shock, infrared, freezing, etc., were employed on a variety of mixed cultures [44, 45] for selective enrichment of acidogenic H₂ producing inoculum. pH also played a critical role in governing the metabolic pathways of the organism where the activity of acidogenic group of bacteria was considered to be crucial. Optimum pH range for the methanogenic bacteria was reported to be between 6.0 and 7.5, while acidogenic bacteria functioned well below the 6 pH [46]. The pH range of 5.5-6.0 was considered to be ideal to avoid both methanogenesis and solventogenesis which was the key for effective H₂ generation. Effect of various pretreatment methods for the production of biohydrogen was outlined in the Table 3. The pretreatment methods used for selective enrichment of hydrogen producing anaerobic consortia considerably influenced the H₂ yield and acid forming pathways of glucose.

Table 3: Comparison of biohydrogen yield by different pretreatment methods

Inoculum	Substrate	Pre-treatment method employed	Max.H ₂ yield(molH ₂ mol glucose ⁻¹)	Optimal pre-treatment method	References
Anaerobi sludge	Sea water culture medium	Acid, alkali, heat-shock, KNO ₃ and control	1.2 ^c	Heat-shock	[47]
Dairy manure wastes	Dairy manure	Acid, NaOH, infrared radiation	31.5 ^b	Acid pretreatment	[48]
Sludge of WTP	Corn Stover	Steam explosion process-neutral and acid	3.0	Steam explosion process with acid	[49]
Digested sludge	Glucose	Acid, base, heat-shock, aeration and chloroform	1.8	Heat-shock	[50]
Cattle manure sludge	Glucose	Freezing and thawing, acid, heat-shock, and sodium 2-bromoethane sulfonate	1.0	Acid	[51]
Cow dung compost	Corn stalk wastes	Acidification	149.69 ^b	Acid	[36]
Methanogenic granules	Glucose	Acid, heat-shock and chloroform	1.2	Chloroform	[52]
Digested wastewater sludge	Sucrose	Heat-shock, aeration, acid, base, 2-bromoethanesulfonic acid and iodopropane	6.12 ^c	Base	[53]
Anaerobic sludge	Dairy wastewater	Sodium 2-bromoethane sulfonate, acid, heat-shock and their combinations	0.0317 ^a	Sodium 2-bromoethanesulfonate	[34]
<i>Clostridium bifermentans</i>	Wastewater sludge	Freezing and thawing, ultrasonication, acidification, sterilization and methanogenic inhibitor	2.1 ^a	Freezing and thawing	[35]
<i>Clostridium bifermentans</i>	Wastewater sludge	Freezing and thawing, sonication, Acidification and	4.1 g kg DS ⁻¹	Freezing and thawing	[54]

<i>Pseudomonas</i> <i>sp.</i> GZ1	Waste sludge	sterilization Sterilization, microwave and ultrasonication	15.02ml g TCOD ⁻¹	Sterilization	[55]
Anaerobic mixed microflora	Glucose	Heat, alkaline and acidification	1405mL	Heat treatment	[56]
Anaerobic mixed culture	Sugar-beet pulp	Alkaline, thermal, microwave, thermal- alkaline and microwave- alkaline	5.15 ^a	Alkaline	[57]
Anaerobic sludge	Corn stover	Microwave assisted acid treatment and thermal acid treatment	1.53	Microwave assisted acid treatment	[58]

a - mmol H₂ g COD⁻¹
b - mLH₂ g TVS⁻¹
c - molH₂ mol sucrose⁻¹

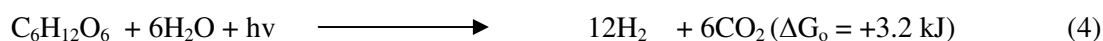
Different pretreatment methods resulted in variations in the fermentation pathways of glucose. Along with an increase in the temperature or use of alkaline or acid pretreatment, the fermentation pathway of glucose converted from ethanol and butyric types to propionate and butyric types, resulting in a decrease in H₂ yield [56]. From the Table 3, after thorough examination, it was strongly recommended that the chemical pretreatment method was most preferable, though few results showed heat shock and base treatment method optimum. Addition of chemicals altered the physiological conditions and environment and enhanced the production and yield. This showed that some inoculum or substrates need suitable treatment based on their waste characteristics. Also there exists some disagreement in the optimum pretreatment method, for example, in the recent published work; Ozkan et al [57] reported alkaline pretreatment on beet pulp is optimum, whereas Liu and Cheng [58] reported microwave-assisted acid pretreatment on corn stover is optimal. Based on these results, it was understood that the possible reason for the disagreement is the difference in their studies with the type of substrates (sugar-beet pulp and corn stover) used for pretreatments, since it has different composition and characteristics in its origin.

2.2.2. Effect of pH

pH is another important factor that influences the activities of hydrogen producing bacteria and the fermentative hydrogen production, because it may affect the metabolism pathway. It has been examined that in an appropriate range, increasing pH could increase the ability of hydrogen producing bacteria to produce hydrogen. Since some studies on fermentative hydrogen production were conducted in batch mode with pH control, while some others were conducted in continuous mode, in these cases, the effect of pH on fermentative hydrogen production was investigated; there exists certain disagreement on the optimal pH for fermentative hydrogen production. For example, the optimal pH for fermentative hydrogen production reported by Mu et al [59] was 4.2, while that reported by Zhao and Yu [60] was 7.0. The possible reason for this disagreement was the difference among these studies in terms of inoculum, substrate and pH range studied. Similarly optimal H₂ production appears to take place with a pH of 5.0 to 6.0 for food wastes [61, 62]; whereas a neutral pH is recommended for crop residues and animal manure [63]. Two different types of experimentation have been performed to determine the optimal pH: one involved adjusting different initial pH in a series of batch tests while the other maintained the same pH in continuous reactors during the fermentation process [63, 64]. In addition, sucrose was the most widely used substrate during the investigation of the effect of pH on fermentative hydrogen production. Thus, investigating the effect of pH on fermentative hydrogen production using organic wastes as substrate is recommended. Ren et al reported the batch experiments with maximum hydrogen yield of 1.77mmol/mmol of glucose were achieved at pH 6.0. Low hydrogen yields observed at pH 4.0 are due likely to inhibitory effects on the microbial growth, although a low pH can be thermodynamically favorable for hydrogen production. Lower yields not only attributed to thermodynamically unfavorable, but also metabolically unfavorable for hydrogen production [65].

2.3. Photo-fermentation

Photo-fermentation differs from dark fermentation because it only proceeds in the presence of light. Hydrogen production by purple non sulfur bacteria was mainly due to the presence of nitrogenase under oxygen-deficient conditions using light energy and reduced compounds (organic acids). Photosynthetic bacteria undergo anoxygenic photosynthesis with organic compounds or reduced sulfur compounds as electron donors. Some non-sulfur photosynthetic bacteria were potent hydrogen producers, utilizing organic acids such as lactic, succinic and butyric acids, or alcohols as electron donors. Since light energy was not required for water oxidation, the efficiency of light energy conversion to hydrogen gas by photosynthetic bacteria was in principle much higher than that by cyanobacteria. Hydrogen production by photosynthetic bacteria was mediated by nitrogenase activity, although hydrogenases might be active for both hydrogen production and hydrogen uptake under some conditions. Photosynthetic bacteria had long been studied for their capacity to produce significant amounts of H_2 . The advantage of their use was in the versatile metabolic capabilities of these organisms and the lack of Photo system II, which automatically eliminates the difficulties associated with O_2 inhibition of H_2 production. These photo-heterotrophic bacteria have been found suitable to convert light energy into H_2 using organic wastes as substrate in batch processes, continuous cultures [66] or immobilized whole cell system using different solid matrices like agar gel and polyurethane foam. The overall reaction of hydrogen production is as follows:



Major drawbacks of the process involved low photochemical efficiencies (3-10 %). This might be overcome by using co-cultures having different light utilization characteristics. Production of biohydrogen by photo fermentation using different substrates and inoculum were outlined in Table 4. It was clear from the Table 4 that, Koku et al. [69] reported the hydrogen production rate of 5.0mLH₂/L h, whereas Basak and Das [70] have reported the maximum production rate of 6.55mLH₂/L h from the same substrate Maleic acid. The possible reason for the deviation in their production rate is on the characteristics of microorganism. It was clearly evident that, production rate and yield could be improved by studying the characteristics and engineering the hydrogenase enzyme to enhance its activity.

Table 4: Comparison of Biohydrogen production rate by various photo-fermentation processes

System	Organism	Substrate	H ₂ production rate, (mL ⁻¹ h ⁻¹)	References
Batch	<i>Rhodopseudomonas</i>	Starch	25.0	[67]
Batch	<i>Rhodobacter capsulata</i>	Starch	0.88	[68]
Batch	<i>Rhodopseudomonas</i>	Acetate	25.2	[67]
Batch	<i>Rhodobacter capsulata</i>	Acetate	0.88	[68]
Batch	<i>Rhodobacter sphaeroides</i> RV	Lactate	62.5	[66]
Batch	<i>Rhodobacter sphaeroides</i>	Malate	5.0	[69]
Continuous	<i>Rhodobacter sphaeroides</i> O.U 001	Malic acid	6.55	[70]
Continuous	<i>Rhodobacter sphaeroides</i> GL-1	Lactate	2100	[71]
Batch	<i>Anabaena variabilis</i>	Water	2.0	[72]

The fermentation process for hydrogen production has been investigated so far, but yet there is lack of details regarding the application of kinetic studies in the photo-fermentation. This review has shown the kinetic studies on biohydrogen production by using the modified Gompertz equation for fitting the

experimental data of cumulative hydrogen production [68, 46, 59, 73]. The modified Gompertz equation is:

$$H(t) = H \exp \left\{ -\exp \left[\frac{R_{\max} \times e}{H} (\lambda - t) + 1 \right] \right\} \quad (5)$$

$H(t)$ is cumulative hydrogen production(g/l) during the fermentation time $t(h)$, H is the maximum gas production potential (g/l), R_{\max} is the maximum production rate(g/lh), λ is lag time to exponential product formed(h) and e is 2.7183. The values of H , R_{\max} and λ are normally determined by best fitting the experimental data using the software curve Expert 1.3[74]

2.4. Two-Stage Fermentation

Combined dark and photo fermentation was a rather new approach in biological hydrogen gas production. It has certain advantages over single stage dark-fermentation or photo-fermentation processes. A three step process scheme consisting of pre-treatment-hydrolysis, dark fermentation and photo-fermentation could be used for this purpose. The first step of pre-treatment includes grinding, acid hydrolysis, neutralization and nutrient balancing to produce carbohydrate solution from the biomass. Fermentable sugars were converted to organic acids, CO_2 and hydrogen in the dark fermentation phase. Light-fermentation was used for production of hydrogen from organic acids under anaerobic conditions in the presence of light. The effluent of dark fermentation in hydrogen production provides sufficient amount of organic acids for the photo-fermentation. Therefore, the limitation by the organic acid availability would be eliminated. Using dark and photo fermentative bioreactors hybrid fermentation technology might be one of the promising routes for the enhancement of H_2 production yields. The synergy of the process lie in the maximum conversion of the substrate which otherwise fail to achieve a complete conversion due to thermodynamic and other limitations [75]. Thus, in this system the light independent bacteria and light dependent bacteria provide an integrated system for maximizing the H_2 yield. Further utilization of organic acids by photo-fermentative bacteria could provide better effluent quality in terms of COD. However, the system should be well controlled to provide optimum media composition and environmental conditions for the two microbial components of the process. In such a system, the anaerobic fermentation of carbohydrate (or organic wastes) produces intermediates, such as low molecular weight organic acids, which are then converted into H_2 by the photosynthetic bacteria in the second step in a photo-bioreactor. Higher hydrogen production yields could be obtained when two systems are combined. In the Table 5, it was observed that maximum yield of 14.2mol/mol sucrose was achieved from sucrose by utilizing *Caldicellulosiruptor* and photosynthetic bacteria of *Rhodospseudomonas capsulatus*.

Table 5: Comparison of Biohydrogen production and yield by two stage dark and photo-fermentation processes

Substrate	Inoculum	H_2 yield (molmol glucose ⁻¹)	Rate of H_2 production (mL ⁻¹ h ⁻¹)	References
Beet molasses	<i>Caldicellulosiruptor</i> and <i>Rhodospseudomonas capsulatus</i>	4.2	7.1	[76]
Sucrose	<i>Clostridium pasteurianum</i> and <i>Rhodospseudomonas palustris</i> WP3-5	14.2mol H_2 mol sucrose ⁻¹		[77]
Starch manufac turing waste	<i>C. butyricum</i> and <i>Enterobacter</i> HO-39 and <i>Rhodobacter</i> sp. M-19	7.2	-	[78]
Cassava starch	<i>Clostridium</i> sp and <i>Rhodospseudomonas</i> <i>palustris</i>	6.07mol H_2 mol sucrose ⁻¹	334.8	[79]
Glucose	<i>Lactobacillus delbrueckii</i> and <i>Rhodobacter sphaeroides</i> RV	7.1		[80]
Olive mill wastewater	Activated sludge and <i>Rhodobacter</i> <i>sphaeroides</i> O.U.001	-	11	[81]
Glucose	<i>E.coli</i> HD701and <i>Rhodobacter</i> <i>sphaeroides</i> O.U.001	-	5.2	[82]
Water hyacinth	Mixed bacteria and immobilized <i>Rhodospseudomonas palustris</i>	596ml H_2 gTVS ⁻¹	-	[83]
Food waste	Anaerobic mesophilic and thermophilic	1.8	-	[84]

Glucose	acidogenesis <i>Enterobacter cloacae</i> DM11 and <i>Rhodobacter sphaeroides</i> O.U.001	5.3	-	[85]
Makkoli wastewater	<i>Clostridium butyricum</i> NCIB 9576 and <i>Rhodobacter sphaeroides</i> E15-1	-	16	[86]
Glucose	<i>C. butyricum</i> and <i>Rhodopseudomonas</i> <i>palustris</i>	5.48	100	[87]

The major challenge in biohydrogen production by dark and light fermentation was to improve the rate and the yield of hydrogen production for an economic process. Biological and engineering studies must be concentrated on these issues. Raw material cost was another concern in biohydrogen fermentations. Therefore, waste materials and renewable resources such as biomass were strongly recommended to utilize as a substrate. Also in the case of two stage fermentation, there exists certain disagreement that Su et al [87] reported a maximum hydrogen yield of 5.48mol/mol glucose, whereas Cheng et al [79] reported a maximum yield of 6.07mol/mol hexose. The possible reason for the disagreement is the difference in their utilization of substrate as glucose and starch. The later was inoculated with immobilized cells and hence its yield was increased.

III. COST COMPARISON OF HYDROGEN PRODUCTION PROCESSES

The cost of H₂ generated from biological processes and other available conventional processes were tabulated in Table 6. Biological hydrogen was comparatively high than that of hydrogen from pyrolysis. Strategies should therefore be developed to effectively lower the cost of H₂ production for commercialization. From the Table 6, it was proven, hydrogen production by conventional methods are not affordable except in the case pyrolysis. Hence, it was shown that despite of its high energy content; the cost of biological hydrogen production was still not a cost effective when compared to the existing pyrolysis method (conventional hydrogen production). It was strongly recommended therefore for the future research in the biological method to overcome or to replace efficiently, the available conventional processes with the cost effective method for biological hydrogen production.

Table 6: Comparison of unit cost of hydrogen production processes with conventional processes

Name of the processes	Raw materials	Energy content of the fuel (MJ kg ⁻¹)	Unit cost of energy content of the fuel US \$/ MBTU ⁻¹
Photo-biological hydrogen	H ₂ O, organic acids	142	10
Fermentative hydrogen	Molasses	-	10
Pyrolysis for hydrogen production	Coal, biomass	-	4
H ₂ from advanced electrolysis	H ₂ O	-	11
H ₂ from Nuclear Energy	Electrolysis and water splitting	-	12-19
H ₂ by biomass gasification	Biomass	-	44-82
H ₂ from Wind Energy	Wind mill	-	34
H ₂ from Photovoltaic power station	Solar energy	-	42
H ₂ from thermal decomposition of steam	H ₂ O	-	13
H ₂ from photochemical	Organic acids	-	21
Gasoline	Crude petroleum	43.1	6
Fermentative ethanol	Molasses	26.9	31.5
Biodiesel	Jatropha seeds	37	0.4
Natural gas	Raw natural gas	33-50	10

Among the alternative energies, at present, biomass and bio-fuel are the ones closer to the parity in conventional and distributed systems, respectively. Increased efforts in the development of advanced

technologies to improve the technical feasibility and scalability of the hydrogen production based on renewable energy, higher carbon emission, large investment growth in renewable energies, etc., could make cost parities to be reached in the near future. Also, as the hydrogen production based in renewable technologies avoid fuel prices in uncertainties, mainly produced in the natural gas market, massive investment in hydrogen production with renewable technologies could be produced before the parities are reached [88].

IV. LIMITATIONS IN BIOLOGICAL HYDROGEN PRODUCTION

Few limitations are somehow discussed here for the smooth transition from the fossil fuel based economy to the hydrogen energy based economy as follows:

- Direct bio-photolysis processes, though essentially attractive, seems to suffer from the intractable barriers of oxygen sensitivity, intrinsic limitations in light conversion efficiencies, and very tedious economics.
- In indirect bio-photolysis, the use of nitrogenase enzyme with its inherent high energy demand and the low solar energy conversion efficiencies are the insurmountable factors.
- Although biological processes for the production of gaseous hydrogen have been well demonstrated with cultured microalgal biomass, these processes must still be integrated into a system capable of meeting basic requirements for overall efficiency of converting solar energy into fuels.
- Processing of some biomass feed stock was too costly and therefore need to develop low cost methods for growing, harvesting, transporting and pretreating energy crops and/or biomass waste products.
- There was no clear contender for a robust, industrially capable microorganism that can be metabolically engineered to produce more hydrogen.
- Several engineering issues need to be addressed which include the appropriate bioreactor design, difficult to sustain steady continuous H_2 production rate in the long term, scale-up, preventing interspecies H_2 transfer in non sterile conditions and separation/purification of H_2 .
- Sensitivity of hydrogenase enzyme to O_2 and H_2 partial pressure severely decreases the efficiency of the processes.
- Insufficient knowledge on the metabolism of H_2 producing bacteria and the levels of H_2 concentration tolerance of these bacteria.
- A lack of understanding on the improvement of economics of the process by combination of H_2 production with other processes.
- The productivity and yield of H_2 from any of the processes explained above was low for commercial application.

To overcome these constraints a numerous improvements need to be done in the future research. The advancements in the scientific research such as development of bioreactor design, engineering of hydrogenase enzyme and genetic modification of microorganism are therefore strongly recommended to improve the yield and rates of production. Many studies are currently carrying out in these technical and scientific advancements for better output as a futuristic goal.

V. CONCLUSIONS

Biohydrogen production was the most challenging area with respect to environmental problems. Due to the energy potential of hydrogen, new processes were to be developed for sustainable hydrogen production in biological methods. This review emphasized the raw material cost as major limitations for bio-hydrogen production showing that the utilization of some carbohydrate rich or starch containing solid wastes or some industry wastewaters are an attractive approach for bio-hydrogen production. Many of the results have proved in better hydrogen yield from different substrates and inoculum followed by their effective and appropriate pretreatment technique. Also it was examined that the sequential or combined bioprocesses of dark and photo-fermentations seem to be the most attractive approach as compared to other methods used for bio-hydrogen production from carbohydrate rich wastes. Two major aspects need indispensable optimization, viz., a suitable substrate and ideal microbial culture that can convert the substrate efficiently to hydrogen. A

comparative study on available processes indicated that biohydrogen production requires greater improvement on the process mainly with respect to hydrogen yield from the cheaper raw materials. Hydrogen production can be improved with the scientific advancements such as genetic modification of organism, engineering of hydrogenase enzyme, by using improved bioreactor and also by hybrid process. The future of biological H₂ production mainly not only depends on research advances, i.e. improvement in efficiency through genetically engineering microorganisms and/or the development of bioreactors, but also on economic considerations (the cost of fossil fuels), social acceptance, and the development of hydrogen energy systems.

ACKNOWLEDGEMENT

This review work has greatly encouraged and supported by Dr. A. Santhiagu, Head, School of Biotechnology, National Institute of Technology Calicut, Kozhikode, India.

REFERENCES

- [1] Armor, J. N., (1999) "The multiple roles for catalysis in the production of H₂", *Applied Catalysis A: General*, 176, 159-176.
- [2] Asada, Y., Tokumoto, M., Aihara, Y., Oku, M., Ishimi, K., Wakayama, T., (2006) "Hydrogen production by co-cultures of *Lactobacillus* and a photosynthetic bacterium, *Rhodobacter sphaeroides* RV", *Int. J. Hydrogen Energy*, 31, 1509-13.
- [3] Baghchehsaree, B., Nakhla, G., Karamanev, D., Argyrios, M., (2010) "Fermentative hydrogen production by diverse Microflora", *Int. J. Hydrogen Energy*, 35, 5021-5027.
- [4] Barbosa, M. J., Rocha, J. M. S., Tramper, J., Wijffels, R. H., (2001) "Acetate as a carbon source for hydrogen production by photosynthetic bacteria", *J. Biotechnol.*, 85, 25-33.
- [5] Basak, N., Das, D., (2007) "Microbial biohydrogen production by *Rhodobacter sphaeroides* O.U.001 in photobioreactor", *Proceedings of World Congress on Eng. and Computer Sci.*, San Francisco, USA.
- [6] Benemann J R, (1996) "Hydrogen biotechnology: Progress and prospects", *Nat Biotechnol.*, 14, 1101-1103.
- [7] Benemann, J. R., Berenson, J. A., Kaplan, N. O., Kamen, M. D., (1973) "Hydrogen evolution by chloroplast-ferredoxin-hydrogenase system", *Proc. Natl. Acad. Sci.*, 70, 2317-2320.
- [8] Benemann, J. R., (1997) "Feasibility analysis of photobiological hydrogen production", *Int. J. Hydrogen Energy*, 22, 979-987.
- [9] Benemann, J. R., (2000) "Hydrogen production by microalgae", *J. Appl. Phycol.*, 12, 291-300.
- [10] Cai, J. L., Wang, G. C., Li, Y. C., Zhu, D. L., Pan, G. H., (2009) "Enrichment and hydrogen production by marine anaerobic hydrogen producing Microflora" *Chinese Sci. Bull.*, 54, 2656-2661.
- [11] Cakir, A., Ozmihci, S., Kargi, F., (2010) "Comparison of biohydrogen production from hydrolyzed wheat starch by mesophilic and thermophilic dark fermentation", *Int. J. Hydrogen Energy*, 35, 13214-13218.
- [12] Chen, C. Y., Yang, M. H., Yeh, K. L., Liu, C. H., Chang, J. S., (2008) "Biohydrogen production using sequential two-stage dark and photo fermentation processes", *Int. J. Hydrogen Energy*, 33, 4755-4762.
- [13] Chen, S. D., Lee, K. S., Lo, Y. C., Chen, W. M., Wu, J. F., Lin, J. Y., Chang, J. S., (2008) "Batch and continuous biohydrogen production from starch hydrolysate by *Clostridium* species" *Int. J. Hydrogen Energy*, 33, 1803-1812.
- [14] Cheng, J., Su, H., Zhou, J., Song, W., Cen, K., (2011) "Hydrogen production by mixed bacteria through dark and photo-fermentation", *Int. J. Hydrogen Energy*, 36, 450-457.
- [15] Cheong, D.Y., Hansen, C.L., (2006) "Bacterial stress enrichment enhances anaerobic hydrogen production in cattle manure sludge", *Appl. Microbiol. Biotechnol.*, 72, 635-643.
- [16] D'Ippolito, G., Dipasquale, L., Vella, F.M., Romano, I., Gambacorta, A., Cutignano, A., Fontana, A., (2010) "Hydrogen metabolism in the extreme thermophile *Thermotoga neapolitana*", *Int. J. Hydrogen Energy*, 35, 2290-2295.
- [17] Das, D., Kanna, N., Veziroglu, T. N., (2008) "Recent developments in biological hydrogen production processes. *Chem. Ind. & Chem. Eng. Quarterly*, 14, 57-67.
- [18] Das, D., Veziroglu, T.N., (2001) "Hydrogen production by biological process: a survey of literature" *Int. J. Hydrogen Energy*, 26, 13-28.
- [19] Datar, R., Huang, J., Maness, P. C., Mohagheghi, A., Czernik, S., Chornet, E., (2007) "Hydrogen production from the fermentation of corn stover biomass pretreated with a steam-explosion process" *Int. J. Hydrogen Energy*, 32, 932 -939.
- [20] Doi, T., Matsumoto, H., Abe, J., Morita, S., (2010) "Application of rice rhizosphere microflora for hydrogen production from apple pomace", *Int. J. Hydrogen Energy*, 35, 7369-7376.

- [21] Doi, T., Matsumoto, H., Abe, J., Morita, S., (2009) "Feasibility study on the application of rhizosphere microflora of rice for the biohydrogen production from wasted bread" *Int. J. Hydrogen Energy*, 34(4), 1735-43.
- [22] Eroglu, E., Eroglu, I., Gunduz, U., Turker, L., Yucel, M., (2006) "Biological hydrogen production from olive mill wastewater with two stage processes", *Int. J. Hydrogen Energy*, 31, 1527-1535.
- [23] Ewan, B C R., Allen, R W K., (2008) "A figure of merit assessment of the routes to hydrogen" *Int. J. Hydrogen Energy*, 30(8), 809-819
- [24] Fan, Y.T., Zhang, Y.H., Zhang, S.F., Hou, H.W., Ren, B.Z., (2006) "Efficient conversion of wheat straw wastes into biohydrogen gas by cow dung compost", *Bioresour. Technol.*, 97, 500-505.
- [25] Fang, H.H.P., Liu, H., (2002) "Effect of pH on hydrogen production from glucose by mixed culture" *Bioresour. Technol.*, 82, 87-93.
- [26] Fang, H.H.P., Liu, H., Zhang, T., (2005) "Phototrophic hydrogen production from acetate and butyrate in wastewater" *Int. J. Hydrogen Energy*, 30, 785-793.
- [27] Fascetti, E., Todini, O., (1995) "*Rhodobacter sphaeroides* RV cultivation and hydrogen production in a one- and two-stage chemostat", *Appl. Microbiol. Biotechnol.*, 22, 300-305.
- [28] Federov, A.S., Tsygankov, A.A., Rao, K.K., Hall, D.O., (1998) "Hydrogen photoproduction by *Rhodobacter sphaeroides* immobilized on polyurethane foam" *Biotechnol. Lett.*, 20, 1007-1009.
- [29] Ferchichi, M., Crabbe, E., Gwang, G.H., Hintz, W., Almadidy, A., (2005) "Influence of initial pH on hydrogen production from cheese whey", *J. Biotechnol.*, 120, 402-409.
- [30] Gadhamshetty, V., Arudchelvam, Y., Nirmalkhandan, N., Johnson, D., (2010) "Modelling dark fermentation for biohydrogen production: ADM1 based model vs Gompertz model", *Int. J. Hydrogen Energy*, 35, 479-490.
- [31] Ghirardi, M. L., Zhang, L., Lee, J. W., Flynn, T., Seibert, M., Greenbaum, E., Melis, A., (2000) "Microalgae: A green source of renewable hydrogen", *TIBTECH*, 18, 506-511.
- [32] Gielen, D. and Simbolotti, G., (2005) "Prospects for hydrogen and fuel cells" *Energy Technology Analysis. International Energy Agency*.
- [33] Guo, L., Li, X.M., Bo, X., Yang, Q., Zeng, G.M., Liao, D.X., (2008) "Impacts of sterilization, microwave and ultrasonication pretreatment on hydrogen producing using waste sludge", *Bioresour. Technol.*, 99, 3651-3658.
- [34] Guo, X.M., Trably, E., Latrille, E., Carrere, H., Steyer, J.P., (2010) "Hydrogen production from agricultural waste by dark fermentation: A Review", *Int. J. Hydrogen Energy*, 35, 10660-673
- [35] Hallenbeck, P.C., Benemann, J.R., (2002) "Biological hydrogen production; fundamental limiting processes. *Int. J. Hydrogen Energy*, 27, 1185-1193.
- [36] Hu, B., Chen, S.L., (2007) "Pretreatment of methanogenic granules for immobilized hydrogen fermentation" *Int. J. Hydrogen Energy*, 32, 3266-3273.
- [37] Idania, V.V., Richard, S., Derek, R., Noemi, R.S., Hector, M.P.V., (2005) "Hydrogen generation via anaerobic fermentation of paper mill wastes", *Bioresour. Technol.*, 96, 1907-13.
- [38] Kanai, T., Imanaka, H., Nakajima, A., Uwamori, K., Omori, Y., Fukui, T., Atomi, H., Imanaka, T., (2005) "Continuous hydrogen production by the hyper-thermophilic archaeon, *Thermococcus kodakaraensis* KOD1", *J. Biotechnol.*, 116, 271-282.
- [39] Kapdan, I.K., Kargi, F., (2006) "Bio-hydrogen production from waste materials", *Enzyme Microb. Technol.*, 38, 569-582.
- [40] Kim, J., Park, C., Kim, T.H., Lee, M., Kim, S., Kim, S.W., Lee, J., (2003) "Effects of various pretreatments for enhanced anaerobic digestion with waste activated sludge", *J. Biosci. Bioeng.*, 95, 271-275.
- [41] Kim, M.S., Lee, T.J., Yoon, Y.S., Lee, I.G., Moon, K.W., (2001) "Hydrogen production from food processing wastewater and sewage sludge by anaerobic dark fermentation combined with photo-fermentation" In: Miyake, J., Matsunaga, T., San Pietro, A. (Ed), *Biohydrogen II*, Elsevier Science: Oxford, 263-272.
- [42] Kim, S., Han, S., Shin, H., (2006) "Effect of substrate concentration on hydrogen production and 16s rDNA-based analysis of the microbial community in a continuous fermentor", *Process Biochem.*, 41(1), 199 - 207.
- [43] Koku, H., Eroglu, I., Gunduz, U., Yucel, M., Turker, L., (2002) "Aspects of metabolism of hydrogen production by *Rhodobacter sphaeroides*", *Int. J. Hydrogen Energy*, 27, 1315-1329.
- [44] Kotay, S.M., Das, D., (2010) "Microbial hydrogen production from sewage sludge bio-augmented with a constructed microbial consortium" *Int. J. Hydrogen Energy*, 35, 10653-59.
- [45] Kotay, S.M., Das, D., (2007) "Microbial hydrogen production with *Bacillus coagulans* IIT-BT S1 isolated from anaerobic sewage sludge" *Bioresour. Technol.*, 98, 1183-1190.
- [46] Kraemer, J.T., Bagley, D.M., (2007) "Improving the yield from fermentative hydrogen production" *Biotechnol. Lett.*, 29, 685-695.

- [47] Kumar, N., Das, D., (2000) "Enhancement of hydrogen production by *Enterobacter cloacae* IIT-BT 08", *Process Biochem.*, 35, 589-593.
- [48] Lee, K.S., Lo, Y.S., Lo, Y.C., Lin, P.J., Chang, J.S., (2004) "Operating strategies for biohydrogen production with high-rate anaerobic granular sludge bed bioreactor", *Enzyme Microb. Technol.*, 35, 605-612.
- [49] Lee, Z., Li, S., Lin, J., Wang, Y., Kuo, P., Cheng, S.S., (2008) "Effect of pH in fermentation of vegetable kitchen wastes on hydrogen production under a thermophilic condition", *Int. J. Hydrogen Energy*, 33(19), 5234 - 41.
- [50] Lemus, R. G., Martinez-Duart, J. M., (2010) "Updated hydrogen production costs and parities for conventional and renewable technologies", *Int. J. Hydrogen Energy*, 35, 3929-3936.
- [51] Lin, C.Y., Chang, R.C., (2004) "Fermentative hydrogen production at ambient temperature", *Int. J. Hydrogen Energy*, 29, 715-720.
- [52] Liu, C. Z., Cheng, X. Y., (2010) "Improved hydrogen production via thermophilic fermentation of corn stover by microwave assisted acid pretreatment", *Int. J. Hydrogen Energy*, 35, 8945-8952.
- [53] Liu, G., Shen, J., (2004) "Effects of culture medium and medium conditions on hydrogen production from starch using anaerobic bacteria" *J. Biosci. Bioeng.*, 98, 251-256.
- [54] Liu, I. C., Whang, L. M., Ren, W. J., Lin, P. Y., (2011) "The effect of pH on the production of biohydrogen by clostridia: Thermodynamic and metabolic considerations", *Int. J. Hydrogen Energy*, 36, 439-449.
- [55] Logan, B.E., (2004) "Feature article: biologically extracting energy from wastewater: Biohydrogen production and microbial fuel cells" *Environ. Sci. Technol.*, 38, 160A-167A.
- [56] Minnan, L., Jinli, H., Xiaobin, W., Huijuan, X., Jinzao, C., Chuannan, L., Fengzhang, Z., Liangshu, X., (2005) "Isolation and characterization of a high H₂ producing strain *Klebsiella oxytoca* HP1 from a hot spring", *Res. in Microbiol.*, 156, 76-81.
- [57] Mu, Y., Yu, H.Q., Wang, Y., (2006) "The role of pH in the fermentative H₂ production from an acidogenic granule-based reactor", *Chemosphere*, 64, 350 -358.
- [58] Nath, K., Das, D., (2006) "Amelioration of biohydrogen production by two-stage fermentation process. *Ind. Biotechnol.*, 2, 44-47.
- [59] Nath, K., Das, D., (2005) "Hydrogen production by *Rhodobacter sphaeroides* strain O.U. 001 using spent media of *Enterobacter cloacae* strain DM11" *Appl. Microbiol. Biotechnol.*, 68, 533-541.
- [60] Nath, K., Muthukumar, M., Kumar, A., Das, D., (2008) "Kinetics of two-stage fermentation process for the production of hydrogen", *Int. J. Hydrogen Energy*, 33, 1195-1203.
- [61] Ozgur, E., Mars, A.E., Peksel, B., Louwerse, A., Yucel, M., Gunduz, U., Claassen, P.A.M., Eroglu, I., (2010) "Biohydrogen production from beet molasses by sequential dark and photo-fermentation", *Int. J. Hydrogen Energy*, 35, 511-517.
- [62] Ozkan, L., Erguder, T.H., Demirer, G.N., (2011) "Effects of pretreatment methods on solubilization of beet-pulp and bio-hydrogen production yield" *Int. J. Hydrogen Energy*, 36, 382-389
- [63] Perera, K.R.J., Ketheesan, B., Gadhamshetty, V., Nimalakhandan, N., (2010) "Fermentative biohydrogen production: Evaluation of net energy gain", *Int. J. Hydrogen Energy*, 35, 12224-12233.
- [64] Redwood, M.D., Macaskie, L.E., (2006) "A two-stage, two-organism process for biohydrogen from glucose. *Int. J. Hydrogen Energy*, 31, 1514-1521.
- [65] Ren, N.Q., Chua, H., Chan, S.Y., Tsang, Y.F., Wang, Y.J., Sin, N., (2007) "Assessing optimal fermentation type for bio-hydrogen production in continuous flow acidogenic reactors", *Bioresour. Technol.*, 98, 1774-1780.
- [66] Shin, H.S., Youn, J.H., Kim, S.H., (2004) "Hydrogen production from food waste in anaerobic mesophilic and thermophilic acidogenesis" *Int. J. Hydrogen Energy*, 29, 1355-63.
- [67] Su, H., Cheng, J., Zhou, J., Song, W., Cen, K., Combination of dark and photo-fermentation to enhance production and energy conversion efficiency. *Int. J. Hydrogen Energy*. 2009, 34:8846-8853.
- [68] Su, H., Cheng, J., Zhou, J., Song, W., Cen, K., (2010) "Hydrogen production from water hyacinth through dark and photo-fermentation" *Int. J. Hydrogen Energy*, 35, 8929-8937.
- [69] Sveshnikov, D.A., Sveshnikova, N.V., Rao, K.K., Hall, D.O., (1997) "Hydrogen metabolism of mutant forms of *Anabaena variabilis* in continuous cultures and under nutritional stress", *FEMS Microbiol. Lett.*, 147, 297 - 301.
- [70] Tang, G., Huang, J., Sun, Z., Tang, Q., Yan, C., Liu, G., (2008) "Biohydrogen production from cattle wastewater by enriched anaerobic mixed consortia: Influence of fermentation temperature and pH", *J. Biosci. Bioeng.*, 106, 80-87.
- [71] Ting, C.H., Lin, K.R., Lee, D.J., Tay, J.H., (2004) "Production of hydrogen and methane from wastewater sludge using anaerobic fermentation" *Water Sci. Technol.*, 50, 223-228.
- [72] Van Ginkel, S.W., Oh, S.E., Logan, B.E., (2005) "Biohydrogen gas production from food processing and domestic wastewaters", *Int. J. Hydrogen Energy*, 30, 1535-1542.

- [73] VenkataMohan, S., Babu, V. L., Sarma, P. N., (2008) "Effect of various pre-treatment methods on anaerobic mixed microflora to enhance biohydrogen production utilizing dairy wastewater as substrate. *Bioresour. Technol.*, 99, 59–67.
- [74] VenkataMohan, S., Mohanakrishna, G., Ramanaiah, S.V., Sarma, P.N., (2008) "Integration of acidogenic and methanogenic processes for simultaneous production of biohydrogen and methane from wastewater treatment" *Int. J. Hydrogen Energy*, 33, 2156–2166.
- [75] VenkataMohan, S., Vijayabhaskar, Y., Sarma, P. N., (2007) "Biohydrogen production from chemical wastewater treatment by selectively enriched anaerobic mixed consortia in biofilm configured reactor operated in periodic discontinuous batch mode" *Water Res.*, 41, 2652-2664.
- [76] Wang, C.C., Chang, C.W., Chu, C.P., Lee, D.J., Chang, V.V., Liao, C.S., (2003) "Using filtrate of waste biosolids to effectively produce bio-hydrogen by anaerobic fermentation", *Water Res.*, 37, 2789–93.
- [77] Wang, J., Wan, W., (2008) "Comparison of different pre-treatment methods for enriching hydrogen-producing cultures from digested sludge", *Int. J. Hydrogen Energy*, 33, 2934–41.
- [78] Wang, Y. Y., Ai, P., Hu, C. X., Zhang, Y. L., (2011) "Effects of various pretreatment methods of anaerobic mixed microflora on biohydrogen production and the fermentation pathway of glucose" *Int. J. Hydrogen Energy*, 36, 390-396
- [79] Wu, S.Y., Hung, C. H., Lin, C. N., Chen, H. W., Lee, A.S., Chang, J. S., (2006) "Fermentative hydrogen production and bacterial community structure in high rate anaerobic bioreactors containing silicone-immobilized and self flocculated sludge", *Biotechnol. Bioeng.*, 93, 934-946.
- [80] Xing, Y., Li, Z., Fan, Y., Huo, H., (2010) "Biohydrogen production from dairy manures with acidification pre-treatment by anaerobic fermentation", *Environ. Sci. Pollut. Res.*, 17, 392-399.
- [81] Yokoi, H., Mori, S., Hirose, J., Hayashi, S., (2002) "Microbial production of hydrogen from starch-manufacturing wastes. *Biomass and Bioenergy*, 22, 389–395.
- [82] Yokoyama, H., Waki, M., Moriya, N., Yasuda, T., Tanaka, Y., Haga, K., (2007) "Effect of fermentation temperature on hydrogen production from cow waste slurry by using anaerobic microflora within the slurry", *Appl. Microbiol. Biotechnol.*, 74(2), 474 - 83.
- [83] Yu, H., Zhu, Z., Hu, W., Zhang, H., (2002) "Hydrogen production from rice winery wastewater in an upflow anaerobic reactor by using mixed anaerobic cultures", *Int. J. Hydrogen Energy*, 27, 1359-1365.
- [84] Zhang, H., Bruns, M.A., Logan, B.E., (2006) "Biological hydrogen production by *Clostridium acetobutylicum* in an unsaturated flow reactor", *Water. Res.*, 40, 728–734.
- [85] Zhang, T., Liu, H., Fang, H.H.P., (2003) "Biohydrogen production from starch in wastewater under thermophilic conditions. *J. Environ. Manag.*, 69, 149-156.
- [86] Zhang, Z. P., Tay, J. H., Show, K. Y., Yan, R., Liang, D. T., Lee, D. J., Jiang, W. J., (2007) "Biohydrogen production in a granular activated carbon anaerobic fluidized bed reactor", *Int. J. Hydrogen Energy*, 32, 185-191.
- [87] Zhao, Q. B., Yu, H. Q., (2008) "Fermentative H₂ product ion in an up-flow anaerobic sludge blanket reactor at various pH values", *Bioresour. Technol.*, 99, 1353–58.
- [88] Zhu, H.G., Bheland, M., (2006) "Evaluation of alternative methods of preparing hydrogen producing seeds from digested wastewater sludge" *Int. J. Hydrogen Energy*, 31, 1980–1988.

Authors Biography

Karthic Pandu: I completed Bachelor of Chemical Technology in Madras University and Master of Chemical Engineering in Annamalai University. I have worked as an Assistant Professor in Periyar Maniammai University from 2007-2009. Currently I am pursuing my research work on biohydrogen production under the supervision of Dr. Shiny Joseph in NITC. My area of research interests are bio-energy and bio-fuels.



Shiny Joseph: I completed my PhD in IIT Madras in 1997. My area of research interests are membrane separation process, bio-energy and bio-fuels and my area of specialization are hydrodynamics and mass transfer studies. I, head the department of chemical engineering from 2006-2010 in National Institute of technology Calicut (NITC) and currently I am working as an Associate professor in the department of chemical engineering, NITC, Kozhikode, Kerala state, India.



MORPHOMETRIC AND HYDROLOGICAL ANALYSIS AND MAPPING FOR WATUT WATERSHED USING REMOTE SENSING AND GIS TECHNIQUES

Babita Pal¹, Sailesh Samanta² and D. K. Pal³
¹ PNG University of Technology Campus

^{2,3} Department of Surveying and Land Studies, PNG University of Technology, Private Mail Bag, Lae- 411, Morobe Province, Papua New Guinea

ABSTRACT

The study analyzes six morphometric parameters namely absolute relief, relative relief, dissection index, average slope, drainage density and ruggedness index, for better understanding of hydrologic processes in a watershed. The advanced application of Remote Sensing (RS) and Geographic Information System (GIS) techniques have lead to estimation of surface runoff and soil loss based on different parameters. Topographical map and Landsat Enhanced Thematic Mapper Plus (ETM+) satellite image are used for morphometric analysis. Land use/land cover, hydrologic soil characteristics, rainfall, curve number (CN) are used for surface runoff assessment using Soil Conservation Service (SCS) model. USLE (Universal Soil Loss Equation) model is used for soil loss estimation with the help of rainfall and runoff factor (R), soil erodibility factor (K), slope length and steepness factor (LS), crop management factor (C) and conservation practice factor (P). These parameters are obtained from monthly and annual rainfall data, soil data, topographic map, satellite image using RS and GIS techniques (with use of Normalized Difference Vegetation Index) respectively. This experimental study is carried out on Watut watershed under Morobe province of Papua New Guinea. The Watut watershed encompasses an area of about 5410.74 sq km. The average drainage density of this watershed is computed at 0.5 km/sq km with the average slope measuring about 31%. The result indicates an average of 68.23% of total rainfall flowing out as surface runoff with a concomitant wearing away of about 6579914 tons/year (12.16 tons/ha/year) of eroded soil in the Watut watershed. Wall to wall (pixel wise) spatial mapping for the entire watershed is carried out using these results. The study underscores that the integrated approach of SCS and USLE model with RS and GIS technologies have great potential for modeling of different hydrological parameters and producing risk maps in any watershed region.

KEYWORDS: Remote Sensing, GIS, Runoff, Soil Loss, Hydrologic and Morphometric Analysis

I. INTRODUCTION

Watershed is a natural laboratory of hydrology. It can be defined as the area that drains the entire precipitation into a particular stream outlet. In other word it is the catchment's area from which all precipitation i.e. rainfall as well as snow melt water drained into a single stream. It forms naturally to dispose the runoff as efficiently as possible. It is a natural convergent mechanism which consists of a network / branch of streamlets converging into a major stream. Studies of morphometry and hydrologic analysis on different watersheds have been carried out in many parts of the world. Relief and climate are the key determinants of running water ecosystems functioning at the basin scale (Lotspeich and Platts 1982, Frisselet al. 1986). Morphometric descriptors represent relatively simple approaches to describe basin processes and to compare basin characteristics (Mesa 2006) and enable an enhanced understanding of the geomorphic history of a drainage basin (Strahler 1964). Drainage

basin morphometric parameters can be used to describe the basin characteristics. These are basin size (stream order, stream length, stream number, and basin area), basin shape factors (circularity ratio, elongation ratio, form factor and compaction ratio), basin perimeter, bifurcation ratios, drainage density, stream frequency and drainage intensity. The risk factor of flood is indirectly related with the bifurcation ratio (Waugh 1996). Quantitative expression of drainage basin shape or outline form was made by Horton (1932) through a form factor. The unit hydrograph, a method for estimating storm runoff, was proposed by Sherman in 1932 and since then it has been considered as a significant concept. Runoff is one of the most important hydrologic variables used in most of the water resources applications. Reliable prediction of quantity and rate of runoff from land surface into streams and river have been difficult and time consuming task especially for un-gauged watersheds. However this information is needed in dealing with many watershed development and management problems (Kumar et.al., 1991). Soil erosion is a complex dynamic process by which productive surface soil is detached, transported and accumulated at a distant place culminating in dual predicament – severe loss of agricultural productivity through the exposure of subsurface soil as well as siltation in reservoirs and natural streams elsewhere (Kandrika and Venkataratnam, 2005). Soil erosion is a major problem throughout the world. Globally, 1964.4 M ha of land is affected by human-induced degradation (UNEP, 1997). Of this, 1,903 M ha are subjected to aggravated soil erosion by water and another 548.3 M ha by wind. The Revised Universal Soil Loss Equation (RUSLE) calculates the long term average annual rate of erosion on a field slope based on rs (Drainage density, average slope, absolute relief, relative relief, dissection index, and ruggedness index mapainfall pattern, soil type, topography, crop system and management practices. RUSLE only predicts the amount of soil loss that results from sheet or rill erosion on a single slope and does not account for additional soil losses that might have occurred from gully, wind or tillage erosion. Economically, the soil loss results in the depletion of arable land and its quality by wearing away the top fertile soil and thereby affecting the land productivity as a whole. The present study is carried out with a view to preparing a land use map of Watut watershed. This will lead to the computation of the runoff volume by developing a suitable hydrological model, thereby enabling us to estimate soil loss by using RUSLE Equation using digital remote sensing data. Finally different hydro-geomorphological maps were developed for morphometric analysis.

II. STUDY AREA AND MATERIALS USED

The Watut watershed is situated in Morobe province. Sixty percent (60%) of the study area is covered by Bulolo districts, twenty one percent area (21%) in Hunon district, fifteen percent (15%) area in Menyamy district and the remaining four percent (4%) in Markham districts. The basin area is bounded within 145° 58' 45.92" E to 146° 53' 22.66" E longitude and 6° 36' 03.09" S to 7° 31' 15.69" S latitude. The basin area (Figure 1) is about 5410.74 sq km. Papua New Guinea's climate is tropical, as one would expect in a country located just south of the Equator. Port Moresby, the capital, and other towns on the coast are quite hot in the summer months; temperatures are considerably cooler in the Highlands. Average temperature is 27 °C while total annual rainfall is 3500 mm in the study area. All other physical characteristics are discussed in the underlying paragraphs.

Topographical maps are used for this study. Optical bands with standard false color combination (SFCC) of LANDSAT-7, Enhance Thematic Mapper Plus (ETM+) satellite images are used to determine the land use/land cover classes in the study area. All other details of the collateral data sets are given in the Table 1.

Table1. List of data used in the study

Sl. No.	Collateral data	Scale/ Resolution	Year of publication	Source
1	Topographical maps	1:250000	1960	National atlas of Papua New Guinea and topographical map
2		1:100000	1973 - 1980	
3	Landsat-7, ETM+	30 m	2001	University of Maryland
4	Soil map	1:500000	1975	USDA
5	Rainfall	Station: point data	1972-2002	weather-forecast.com report

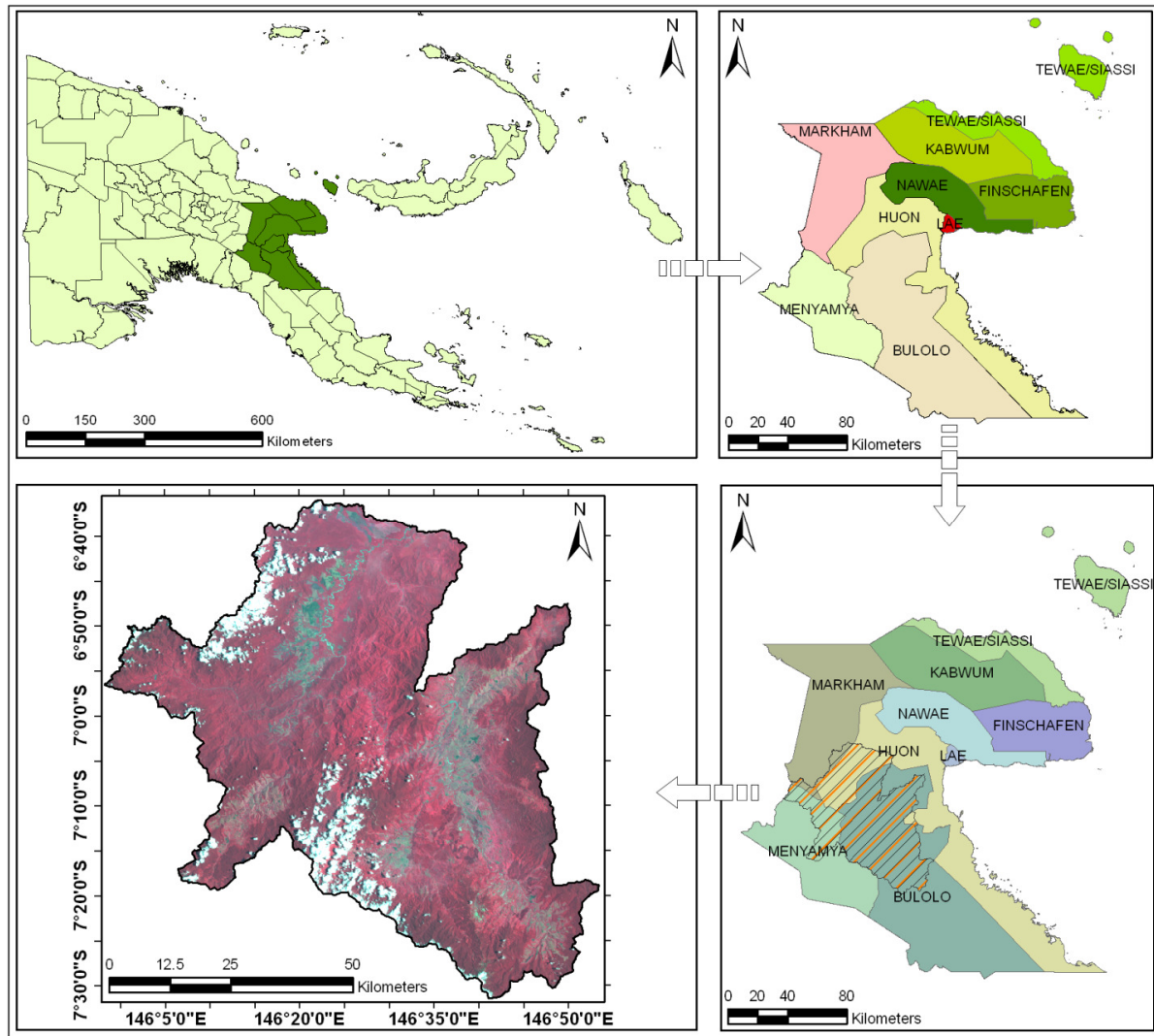


Figure 1. Location map of the Watut watershed region

III. METHODOLOGY

ArcGIS 10.0 and Erdas Imagine software are used for the preparation of different digital data set using satellite image, topographical map and collateral information. Six morphometric analyses, namely absolute relief, relative relief, dissection index, average slope, drainage density and ruggedness index and two hydrological analyses viz. surface runoff and soil loss estimation are performed using different algorithms in the model maker.

3.1. Preparation of drainage, land use/land cover and soil data set

A drainage map is prepared using the National atlas of Papua New Guinea, topographical map and satellite image for delineation of watershed along with the contour information. The drainage network of the entire Watut watershed is subdivided into individual lengths of channel and arranged according to the magnitude of orders as proposed by Strahler (1952). Using the magnitude of order attributes, stream order map is prepared (Figure 2a). The flow direction of the prepared drainage map of Watut watershed is from south-east to north-east. Bulolo river from left and Snake river from right hand side meet each other at $146^{\circ} 35' 29.23''$ E and $7^{\circ} 03' 20.32''$ S. Finally the flow, Watut river meets with Markham River, in the Huon district. Total length of the river is 224.45 km. Other tributaries of this river are Waime, Wafi river, Banir river, Isimp river, Mumena river and Laloona river. The maximum stream order of this drainage network is five which appear in lower part of the basin.

3.1.1 Land use/ land cover

Land use/land cover is one principal tool for runoff and soil loss estimation. The land use /land cover data sets are generated from the digital image classification of LANDSAT, ETM+ satellite images. This classification is performed taking nine classes within the entire study area, namely water body, dense forest, low dense forest, agriculture, shrub land, barren land, plantation, mangrove and settlement (Figure 2b). 80% of the study area comprises forests. In the lower catchment area some agricultural lands are found, albeit of a tiny proportion featuring only about 2% of the study catchment. Detailed land use and land cover statistics are given in Table 2. Overall accuracy achieved is 89%, after carrying out an accuracy assessment using ground truth (reference sample points) data sets.

3.1.2 Soil characteristics

Soil texture and hydrologic soil group are very important parameters for estimation of runoff, soil loss, transport capacity and net detachment. Soil texture is the principal factor affecting soil erodibility factor. Seven types of soil texture i.e. sandy clay, silty clay loam, silty clay, loam, sandy loam, silty loam and sandy clay loam are found in the study area (Figure 2c). The hydrologic soil groups indicate the infiltration potential (vertical flow) of the soil after prolonged wetting. The SCS model has classified all soils into four hydrologic groups. These are (i) "Group A" soils which have low run off potential and high infiltration rates even when thoroughly wetted. They consist chiefly of deep, well to excessively drained sands or gravels while loamy sands or sandy loams are sometimes included in this group where the 'saturated infiltration rate' is greater than 0.76 cm/hr when thoroughly wet. (ii) "Group B" soils that have moderate infiltration rates when thoroughly wetted and consist chiefly of moderately well to well-drained soils with moderately fine to moderately coarse textures. Silt loams or loamy soils embrace this group when the saturated infiltration rate is 0.38 to 0.76 cm/hr. (iii) "Group C" soils have low infiltration rates when thoroughly wetted and consist chiefly of soils with a layer that impedes downward movement of water and also soils with moderately fine to fine texture. Sandy clay loams feature in this group and the saturated infiltration rate is 0.13 to 0.38 cm/hr. (iv) "Group D" soils have high runoff potential. They have very low infiltration rates when thoroughly wetted and consist mainly of clay soils with a high swelling potential, soils with a permanent high water table, soils with a clay pan or clay layer at or near the surface and shallow soils over a nearly impervious material. Clay loams, silt clay loams, sandy clays and silty clays or clays are covered in this group when saturated infiltration rate falls to 0 - 0.13 cm/hr. All four types of hydrologic soil group are found in this watershed region (Figure 2d and table 2).

Table 2. Land use/land cover and hydrological soil group statistics of the study area.

Land use/Land cover	Area (sq km)	Area (%)	Hydrological soil group	Area (sq km)	Area (%)
Dense forest	2848.91	52.7	Group-A	127.52	2.4
Shrub	47.51	0.9			
Low dense forest	1707.89	31.6	Group-B	239.41	4.4
Mangrove	109.25	2.0			
Settlement	1.89	0.0	Group-C	1419.17	26.2
Barren land	638.51	11.8			
Plantation	2.18	0.0	Group-D	3624.64	67.0
Agriculture	14.81	0.3			
Water	39.80	0.7	Total	5410.74	100.0
Total	5410.74	100.0			

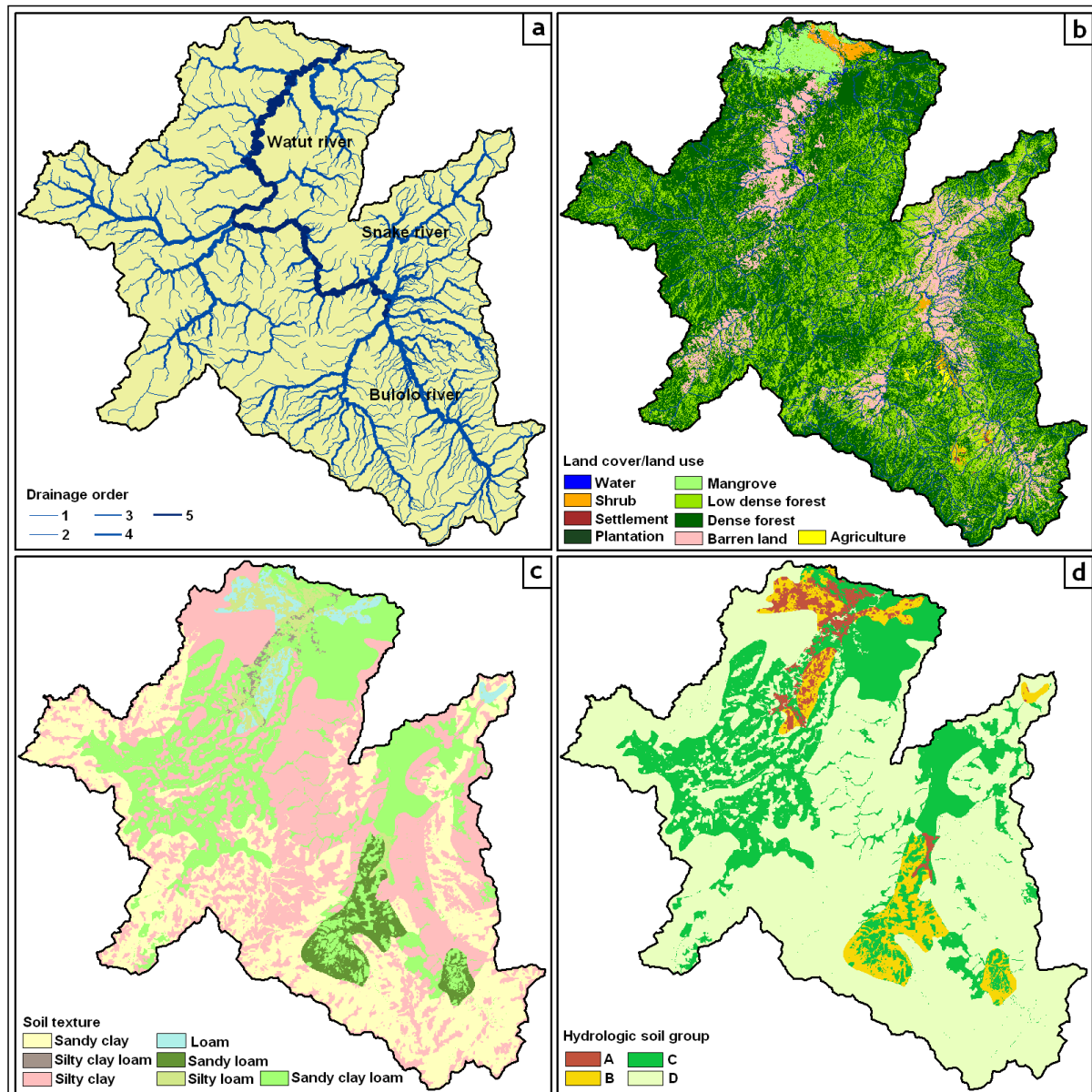


Figure 2. Drainage order (2a), land use/land cover (2b), soil texture (2c) and hydrologic soil group (2d) of Watut watershed, generated from satellite image and soil data.

3.2. Morphometric analysis of the watershed

Morphometry is the measurement and mathematical analysis of the configuration of the earth's surface, shape and dimension of its landforms (Agarwal, 1998; Obi Reddy et al., 2002). A major emphasis in geomorphology over the past several decades has been on the development of quantitative physiographic methods to describe the evolution and behavior of surface drainage networks (Horton, 1945; Leopold & Maddock, 1953; Abrahams, 1984). Different morphometric techniques, like absolute relief, relative relief, dissection index, average slope, drainage density and ruggedness index (Table 3) are considered for quantitative analysis of different attribute of the watershed. A point layer is generated in the ArcGIS s/w environment and the spot heights are collected according to the longitudinal and latitudinal value. Then spot height attribute for each sample location is introduced to the point layer. Arc coverage, which represents the contours, is used to build 3D surfacing in the ERDAS IMAGINE s/w environment. All point data are interpolated to the raster through krigging interpolation (Matheron, 1970) techniques in the ArcGIS s/w environment. Both interpolated and surfacing outputs are merged to create raster digital elevation model (DEM) for the Watut watershed and the absolute relief map is generated as shown in figure 3a.

Table 3. Different morphometric techniques used in the study

Sl. No.	Morphometric techniques	Key description	Data used	Software used
1	Absolute relief	Maximum height of relief	Spot-heights and contours with z values	Erdas Imagine (Surfacing)
2	Relative relief	Maximum – minimum relief	Digital Elevation Model	ArcGIS (Range)
3	Dissection index	Relative relief/Maximum height	Relative relief and Digital Elevation Model	Erdas Imagine (Model maker)
4	Drainage density	length of streams/area	Drainage map	ArcGIS (Line density)
5	Ruggedness index	(Drainage density x relative relief)/1000	Drainage density and relative relief	Erdas Imagine (Model maker)
6	Slope	(Rise/run) x 100	Digital Elevation Model	ArcGIS & Erdas Imagine (3D analysis)

The maximum altitude is 3316 m, demarcated in the southern part and the minimum elevation is closer to mean sea level (MSL) observed in the northern part of the watershed.

3.2.1 Relative relief

To show spatial variation from one place to another calculation of relative relief (Rao, 2011) is paramount. The value of relative relief is calculated from the contour and DEM data using range analysis in ArcGIS software. Figure 3b displays the relative relief characteristics, where maximum relative relief value is 631 m in the middle and eastern parts of the Watut watershed.

3.2.2 Dissection index

For better understanding of morphometry as well as physiographic attribute, dissection index analysis is performed (Schumm, 1956). Dissection index is defined as a ratio between actual dissection made by the rivers and potential dissection up to base levels. Relative relief and maximum altitude are used to compute the dissection index. Higher the value of dissection index, larger is the undulation and instability of the terrain. Maximum dissection index is calculated as 0.58 in the northern part of the watershed (Figure 3c).

3.2.3 Drainage density

The drainage density is the ratio of the total length of streams within a watershed to the total area of the watershed (Horton, 1932). A high value of the drainage density indicates a relatively high density of streams inducing a rapid storm response giving rise to a higher runoff yield. Using the ArcGIS kernel density function the drainage density map of the watershed is prepared (Figure 3d). Maximum drainage density is computed as 1.31 km per sq km in the south and south-east parts of the study area.

3.2.4 Ruggedness index

The topographic ruggedness index too indicates the extent of instability of land surface (Strahler, 1956). Ruggedness index map is generated using relative relief and drainage density data set. Here the values of ruggedness almost coincide with other morphometric values. Ruggedness index of the area varied from 0.01 to 0.45 (Figure 3e). The values of ruggedness index increase towards south and south-east in the watershed.

3.2.5 Average slope

The slope tool calculates the maximum rate of change between each cell and its neighbors. The lower slope value indicates the flatness of the terrain and higher slope value point to a steeper terrain. The output slope raster can be calculated as percent of slope or degree of slope. When the slope angle equals 45 degrees, the vertical rise is equal to the horizontal run and is expressed as 100 percent slope. As the slope approaches vertical (90 degrees), the percentage slope approaches infinity. Slope analysis is performed to prepare the slope data set of the area. Slope of the area is calculated in percent

gradient using ArcGIS 3d analysis tool. The watershed region possesses a maximum slope of 153 percent in the north-west and middle part (Figure 3f).

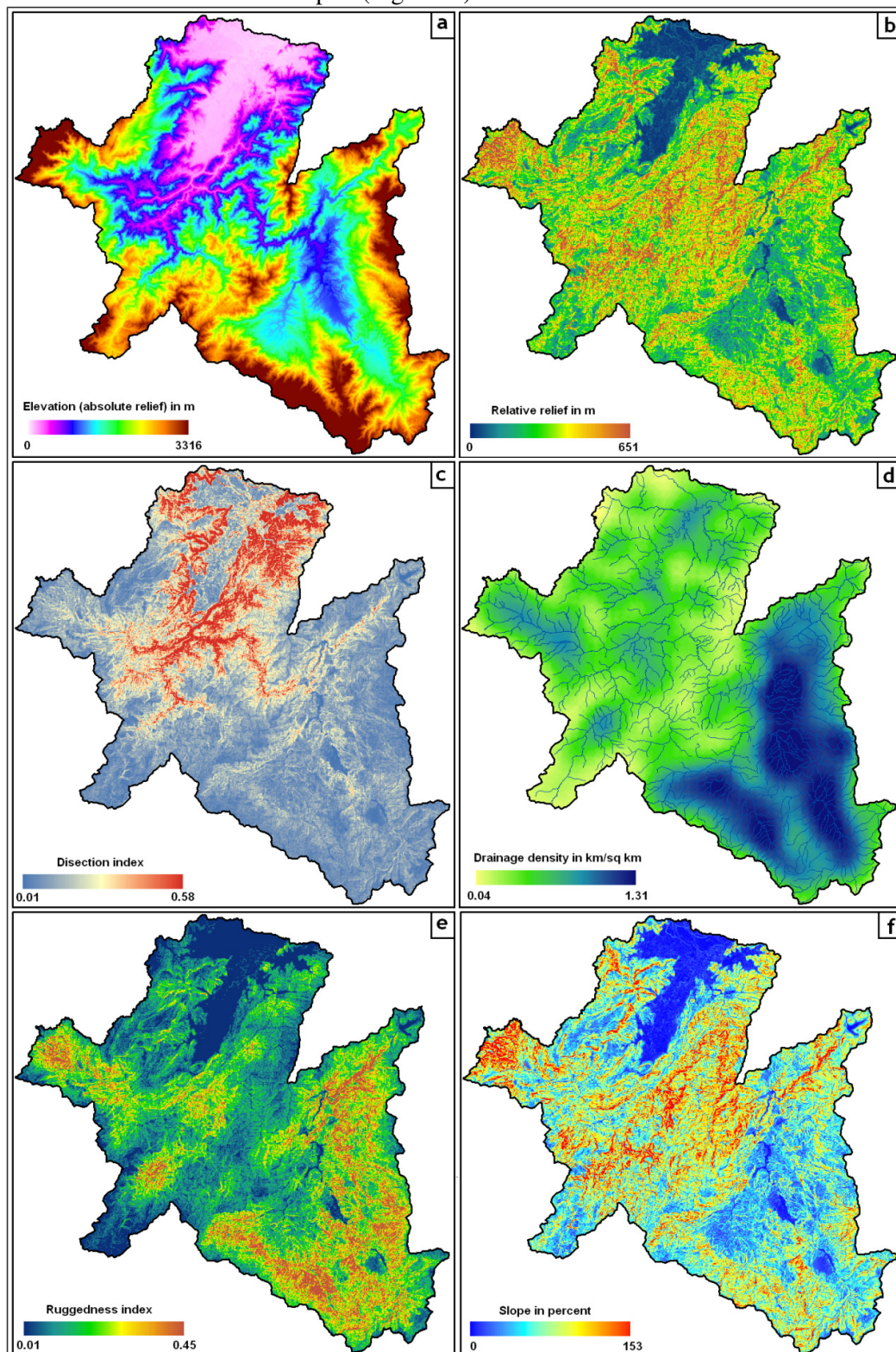


Figure 3. Absolute relief (3a), relative relief (3b), dissection index (3c), drainage density (3d), ruggedness index (3e) and average slope (3f) generated from satellite image and DEM

3.3. Spatial modeling of surface runoff

The SCS method has been used by many researchers because it gives consistently usable results (Rao et al., 1996; Sharma et al. 2001; Chandramohan and Durbude, 2001; Sharma and Kumar, 2002) for

runoff estimation. The SCS model computes direct runoff through an empirical equation that required the rainfall and a watershed co-efficient as input. The watershed co-efficient is called the curve number (CN) which represents the runoff potential of the hydrologic soil cover complexes. The SCS model (SCS, 1972) involves relationships between land cover, hydrologic soil classes and curve number. The equation 1 is used to calculate the surface runoff of the Watut watershed.

$$Q = (P - I_a)^2 / (P - I_a + S) \text{ ----- (1)}$$

Where,

Q is actual surface runoff in mm,

P is rainfall in mm,

I_a is 0.4S [Initial abstraction (mm) or losses of water before runoff begins by soil and vegetation (such as infiltration, or rainfall interception by vegetation)],

S is the potential maximum retention in mm and is calculated using the equation 2.

$$S = (25400 / CN) - 254 \text{ ----- (2)}$$

Where,

CN is curve number of hydrologic soil cover complex, which is a function of soil type, land cover and antecedent moisture condition (AMC), is shown in table 4.

Table 4. Curve numbers for different kinds of land use/land cover and soil (AMC II & I_a=0.4S)

Sl. No.	Land use/land cover	Hydrologic soil group	Curve number	Area in sq km
1	Water	A	100	6.41
		B	100	0.00
		C	100	12.29
		D	100	21.11
2	Dense forest	A	30	65.12
		B	55	106.28
		C	70	702.18
		D	77	1975.33
3	Low dense forest	A	45	15.09
		B	66	93.50
		C	77	401.25
		D	83	1198.07
4	Agriculture	A	62	0.00
		B	71	3.63
		C	78	3.63
		D	81	7.55
5	Shrub land	A	39	20.21
		B	61	3.55
		C	74	9.43
		D	80	14.33
6	Barren land	A	68	20.69
		B	79	32.45
		C	86	246.12
		D	89	339.23
7	Settlement	A	57	0.00
		B	72	0.00
		C	81	1.89
		D	86	0.00
8	Plantation	A	45	0.00
		B	66	0.00
		C	77	0.00
		D	83	2.18
9	Mangrove	A	0	0.00
		B	0	0.00
		C	70	42.38
		D	77	66.84
Total area in sq km				5410.74
Weighted curve number				77

3.4. Spatial modeling of soil loss

Hydrologic calculation of soil loss for Watut watershed is performed using simple equations. By using the revised universal soil loss equation (Renard, 1997), the total soil loss is measured for the Watut watershed. The revised universal soil loss equation (RUSLE) is-

$$A = R \times K \times LS \times C \times P \dots\dots\dots (3)$$

Where,

A is the potential long term average annual soil loss in tons per acre per year. This is the amount, which is compared to the "tolerable soil loss" limits.

R is the rainfall and runoff factor by geographic location, is calculated using the following equation.

$$R = 38.5 + 0.35 \times Pr \dots\dots\dots (4)$$

Where, Pr is average annual precipitation of the study area

K is the soil erodibility factor. K factor is calculated according to the soil texture type of the area (Robert, 2000), e.g. 0.24 for sandy clay, 0.32 for silty clay loam, 0.26 for silty clay, 0.20 for sandy clay loam, 0.30 for loam, 0.13 for sandy loam and 0.38 for silty loam.

LS is the slope length-gradient factor. The following equation (Robert, 2000) is used to calculate the slope length-gradient factor.

$$LS = [0.065 + 0.0456(\text{slope}) + 0.006541(\text{slope})^2] \times (\text{slope length} \div \text{const})^{NN} \dots\dots\dots (5)$$

Where, slope is the percent of steepness (%), slope length is length of slope (m), constant is 22.1 and - NN is 0.5 as slope is more than 1%.

C is the crop/vegetation and management factor, is obtained multiplying crop type factor and the tillage method factor, e.g. 0 for water, 0.002 for settlement, 0.004 for forest area, 0.050 for shrub land and grass cover land, 0.125 for agriculture and 0.5 for barren land.

P is the support practice factor, considered as according to the up and down Slope of the area, e.g. value 0.60 for 0-7% , 0.70 for 7-14% , 0.80 for 14-21% , 0.90 for 21-28% and 1.0 for more than 28% of slope.

IV. RESULT AND DISCUSSION

In upper watershed region number of stream is greater and steep slopes are found. Thus the overland flow velocity is also more, about 5.0 km per hour as measured in the upper watershed region. Relative relief, dissection index, average slope, ruggedness index, drainage density are more, indicating unstable topography and greater vulnerability for erosion and land slide. Lower reaches of the watershed region have gentle slope and fewer number of channel/streams. The dissection index and ruggedness index are also very low. Erosional materials are transported and deposited in this lower part of the Watut river, bereft of the required kinetic energy to carry forward. So the river thalwegs get continually escalated and water holding capacity goes down engendering frequent flash floods. The average bifurcation ratio is also measured at 1.74 for Watut watershed.

The runoff curve numbers for land cover delineation are used for determination of curve number for the watershed. The weighted curve number for AMC II is found 77 for the total watershed area. Similarly the weighted curve numbers for antecedent moisture conditions I & III are also calculated which are 61 and 89 respectively. Calculated curve numbers are used to model surface runoff of the area. The spatial distributions of calculated curve number (Figure 4a) are mapped for each pixel in the raster data format. Model maker tool of Erdas Imagine and spatial analysis tool of ArcGIS are used for this mapping purpose. The average slope-length gradient factor of the watershed is calculated as 20.30 from digital elevation model in the Erdas Imagine model maker. Slope-length gradient factor is mapped pixel-by-pixel for the area as shown in the figure 4b.

Land use/land cover, hydrologic soil characteristics, rainfall, curve number are used to estimate surface runoff by SCS model for Watut watershed. There is a big amount of storm rainfall recorded as 277 mm, in the month on March, 2011. The above rainfall (277 mm) is introduced as an input in the SCS model. As the model output (Figure 4c) the minimum range of surface runoff (125 to 150 mm) is found in the lower watershed area and maximum range (225 to 245 mm) in some parts of the upper watershed area. The calculated average surface runoff is 189 mm, almost 68 percent of the total storm rainfall.

Table 5. Estimation of runoff volume for Watut watershed

Month/date	Rainfall (mm)	AMC	CN	S	Runoff (mm)
15-20 th September	171	II	77	75.87	91.37
January	275				186.74
February	260				172.62
March	277*				188.63*
April	262				174.50
May	195				112.71
June	153				75.78
July	144				68.15
August	117				46.20
September	113				43.09
October	158				80.07
November	179				98.42
December	255				167.94

* Spatial mapping is done using storm rainfall as shown in figure 4C (output: input \rightarrow 1:1.47).

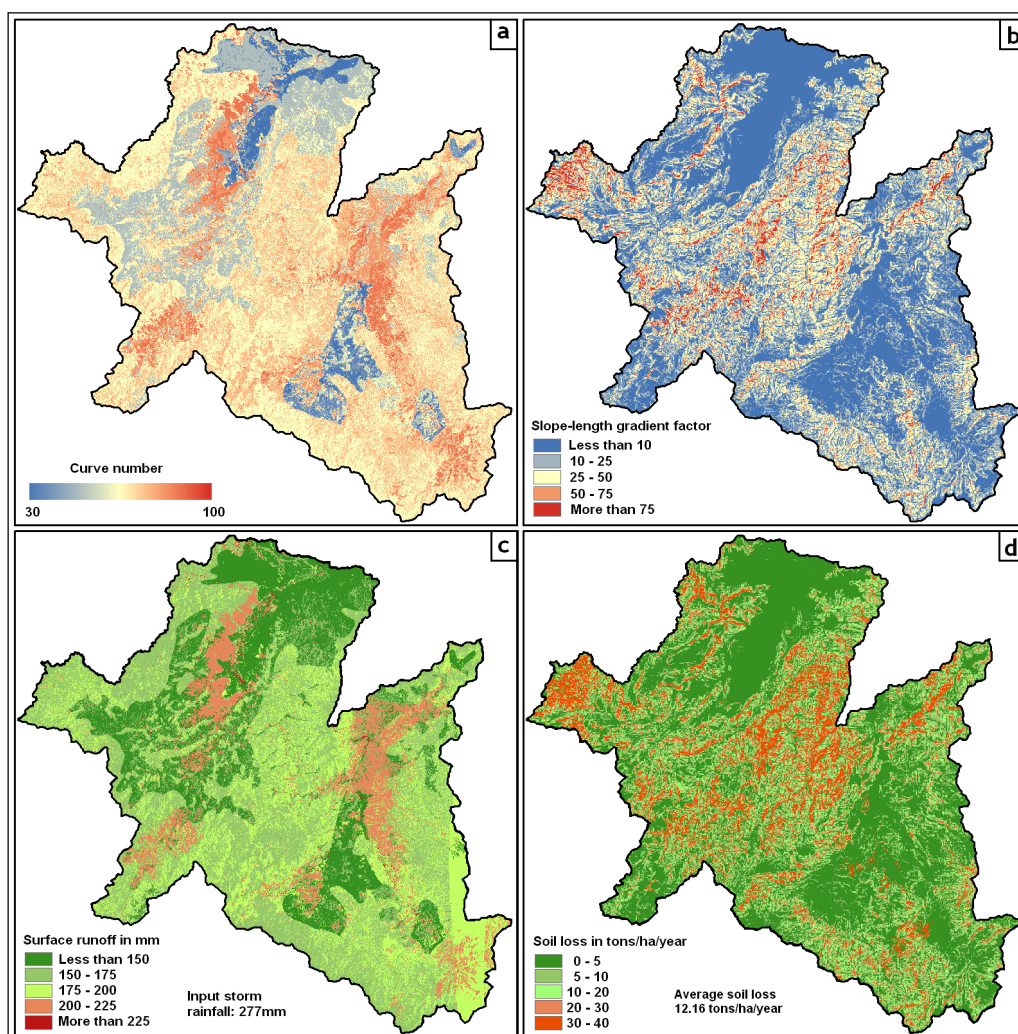


Figure 4. Calculated curve number from land use and hydrological soil group (4a), slope-length gradient factor from DEM (4b), estimated surface runoff (4c) and soil loss (4d) according to SCS and RUSLE model using Erdas Imagine and ArcGIS.

Rainfall and runoff factor, soil erodibility factor, slope length and steepness factor, crop management factor and conservation practice factor are used to calculate the soil erosion in the watershed area following USLE model. Monthly and annual rainfall data, soil map of the region, DEM, RS techniques (with use of normalized difference vegetation index) and land use/land cover map are used

to obtain above parameters. Total soil loss is 6579914 tons/year calculated for the Watut watershed. Soil losses characteristics of the watershed area are shown in figure 4d. Red color pixels indicate high rate of soil erosion in the watershed area. The average soil loss of the area is calculated as 12.16 tons/ha/year.

V. CONCLUSIONS

Watershed is a basic unit for morphometric analysis. Remote sensing and GIS techniques are known for providing very high accuracy in mapping and measurement done in morphometric analysis. Absolute reliefs, relative relief, dissection index, ruggedness index, slope and drainage density are the most useful criteria for the morphometric classification of a watershed. These decisive factors are more in the upper part of the watershed and are instrumental in controlling the runoff pattern, sediment yield and other hydrological parameters of the watershed region. Surface runoff is calculated as very high (more than 200 mm, when the storm rainfall is 277 mm) in some parts of the watershed where land cover types are 'barren land' and / or 'open hills area'. The watershed region has considerably high rate of soil erosion (6579914 tons/year), because of high slope-length gradient factor. In this study, the methodology for determination of runoff, soil loss and spatial mapping are attempted for Watut watershed using RS, GIS, RUSLE and SCS model. This approach could be applied in other watersheds or river basin for planning of various conservation measures.

ACKNOWLEDGEMENTS

One of the authors (SS) expresses sincere gratitude to Papua New Guinea University of Technology & Department of Surveying and Land studies for providing digital image interpretation laboratory facility to carry out this work. The authors are also grateful to the all the academic staff of GIS section for their valuable comments and suggestions.

REFERENCES

- [1] Abrahams, A. D. (1984). Channel networks: a geomorphological perspective. *Water Resour Res.*, 20:161-168
- [2] Agarwal, C. S. (1998) Study of drainage pattern through aerial data in Naugarh area of Varanasi district, U.P., *Jour. Indian Soc. Remote Sensing*, 26, 169-175.
- [3] Bates, N. (1981). Practical foundations of physical geography, Valley shapes. In B. Knap (Ed.), London: George Allen & Unwin, 25-29.
- [4] Chandrmohan, T. and Durbude, D.G. (2001), Estimation of runoff using small watershed models, *Hydrology Journal*, 24 (2), 45-53.
- [5] Chorley, R. J. (1969). The drainage basin as the fundamental geomorphic unit, in *Water, earth, and man: a synthesis of hydrology, geomorphology and socio-economic geography*, R. J. Chorley (Ed.), London: Methuen, 77-99.
- [6] Frissell, C. A., W. J. Liss, C. E. Warren, and Hurley, M. D. (1986). A hierarchical framework for stream habitat classification: viewing streams in a watershed context, *Environmental Management*, 10, 199-214.
- [7] *Handbook of Hydrology* (1972). Soil Conservation Department, Ministry of Agriculture, New Delhi.
- [8] Horton, R. E. (1932). Drainage basin characteristics. *Trans. Am. Geophys. Unions*. 13, 350-361.
- [9] Horton, R. E. (1945). Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, *Geol. Soc. Am. Bull.* 56: 275-370.
- [10] Kandrika, S., and Venkataratnam, L. (2005). A Spatially distributed event based model to predict sediment yield, *Journal of Spatial Hydrology* Spring, 5 (1), 1-19.
- [11] Kumar, P., Tiwari, K. N., and Pal, D. K. (1991). Establishing SCS runoff curve number from IRS digital data Base, *Journal of the Indian Society of Remote Sensing*, 19 (4), 24-251.
- [12] Leopold, L. B. and Maddock, T. (1953). The hydraulic geometry of stream channels and some physiographic implications USGS professional paper 252, pp.1-57.
- [13] Lotspeich, F. B., and Platts, W. S. (1982). An integrated land-aquatic classification system, *North American Journal of Fisheries Management*, 2, 138-149.
- [14] Matheron, G. (1970). The theory of regionalized variables and its applications, Issue 5, *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau*, Paris: École Nationale Supérieure des Mines, pp. 212.

- [15] Mesa, L.M., (2006). Morphometric analysis of a subtropical Andean basin (Tucuman, Argentina), *Environmental Geology*, 50(8), 1235-1242.
- [16] Nayak, T. R., and Jaiswaly, R. K. (2003). Rainfall run off modeling using satellite data and GIS for Beas river in Madhya Pradesh, *IE(I) J.*, 84, 47-50.
- [17] Obi Reddy, G. E., Maji, A. K. and Gajbhiye, K. S. (2002). GIS for morphometric analysis of drainage basins, *GIS India* (11), 4: 9-14.
- [18] Pidwirny, J. (2006). Fundamentals of physical geography, <http://www.fundamentalsofphysicalgeography.net>
- [19] Rao, L. A. K., Ansari, Z. R, and Yusuf, A (2011). Morphometric Analysis of Drainage Basin Using Remote Sensing and GIS Techniques: A Case Study of Etmadpur Tehsil, Agra District, U.P., *International Journal of Research in Chemistry and Environment*, Vol. 1 (2):36-45.
- [20] Rao, K. V., Bhattacharya, A. K. and Mishra, K. (1996). Runoff estimation by curve number method- case studies, *Journal of Soil and Water Conservation*, 40, 1-7.
- [21] Renard, K. G., Forster, G., Weesies, G., McCool, D., and Yoder, D. (1997). Predicting Soil Erosion by Water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE), *Agricultural Handbook*, (Washington DC: USDA), 703.
- [22] Robert, P. S. (2000). Engineer, Soil Management/OMAFRA; Don Hilborn-Engineer, Byproduct Management/OMAFRA.
- [23] Rodda, J. C. (1969). The flood hydrograph, in *Water, earth and man: a synthesis of hydrology, geomorphology and socio-economic Geography*, R. J. Chorley (Ed.), London: Methuen, 405-418.
- [24] Schumm, S.A. (1956). Evolution of drainage system and slope in badlands of Perth Amboy, New Jersey. *Bull. Geol. Soc. Am.*, 67: 597.
- [25] Sharma, D. and Kumar, V. (2002). Application of SCS model with GIS data base for estimation of runoff in an arid watershed, *Journal of Soil and Water Conservation*, 30 (2), 141-145.
- [26] Sherman, L. K. (1932). Stream flow from Rainfall by the Unit-Graph Method, *Eng. News Record*, 08, 501-505.
- [27] Strahler, A. N. (1952). Hypsometric (area-attitude) analysis of erosional topography. *Geol. Soc. Am. Bull.* 63, 1117-1142.
- [28] Strahler, A.N. (1956). Quantitative slope, analysis. *Bull. Geol. Societ. Am.* 67: 571-596
- [29] Strahler, A. N. (1964). Quantitative geomorphology of drainage basins and channel networks. In V. T. Chow (Ed.), *Handbook of Applied Hydrology*, New York, McGraw Hill, 4, 39-76.
- [30] UNEP (1997). *World Atlas of Desertification*, 2nd edition, Arnold London.
- [31] Waugh, D. (1996). *Geography: an integrated approach*, New York.

Authors Profile

Babita Pal has received her M. Sc. degree from Vidyasagar University, India in 2007. Her major research Interests are in Remote Sensing, GIS and Natural Resource Management.



Sailesh Samanta is currently working as a Lecturer in Department of Surveying and Land Studies, Papua New Guinea University of Technology, Morobe Province, Papua New Guinea. He is having overall teaching experience of 6 years. His major research Interests are in Remote Sensing, GIS, Natural Resource Management and climate.



D. K. Pal has received his M Sc Ph.D degree from IIT, KGP, India. He is currently working as a Professor & HOD in Department of Surveying and Land Studies, Papua New Guinea University of Technology, Morobe Province, Papua New Guinea. He was formerly Head – Department of Geography & Environment Management; Founder Coordinator & Head – Department of Remote Sensing & GIS Vidyasagar University, Midnapore, West Bengal, India. He was Senior Scientist, Indian Space Research Organization, Govt. of India. He has over 30 years in the field of Resources management, Land use Studies, Sustainable Development, Disaster Management, Agricultural Chemistry, Soil Science, application of Remote Sensing and GIS in various natural resources.



ADAPTIVE HYSTERESIS BAND CURRENT CONTROL FOR TRANSFORMERLESS SINGLE-PHASE PV INVERTERS

B. Nagaraju¹, K. Prakash²

¹ Assistant Professor, Vaagdevi College of Engineering, Warangal-India

² Associate Professor, Vaagdevi College of Engineering, Warangal-India

ABSTRACT

Current control based on hysteresis algorithms are widely used in different applications, such as motion control, active filtering or active/reactive power delivery control in distributed generation systems. The hysteresis current control provides to the system a fast and robust dynamic response, and requires a simple implementation in standard digital signal platforms. On the other hand, the main drawback of classical hysteresis current control lies in the fact that the switching frequency is variable, as the hysteresis band is fixed. In this paper a variable band hysteresis control algorithm will be presented. As it will be shown, this variable band permits overcoming the aforementioned problem giving rise to an almost constant switching frequency. The performance of this algorithm, together with classical hysteresis controls and proportional resonant (PR) controllers, has been evaluated in three different single-phase PV inverter topologies, by means of simulations performed with PSIM. In addition, the behavior of the thermal losses when using each control structure in such converters has been studied as well.

I. INTRODUCTION

The classical current control techniques for power converters, based on PWM modulation, present a certain control delay that does not permit to control accurately the instantaneous-peak of the currents that are being delivered by a power converter [1]. In order to solve this issue, asynchronous sine-triangle PWM voltage control can be used in conjunction with instantaneous current feedback control. However, when this method is used the error signals are nearly sinusoidal and the dynamical response of the system is limited by the stability bandwidth of the feedback loops. The use of predictive voltage control techniques [2] permits improving the bandwidth of the controller, as the tracking of the reference does not depends exclusively on the controller tuning, but also on the feed-forward control signal provided by the predictive algorithm. Although these techniques tend to give optimum performance in terms of response and accuracy, they depend highly on the accuracy of the predictive model and they are somehow complicated to implement in standard Digital Signal Processors (DSP) [3,4]. As an alternative to PWM techniques hysteresis current controllers have been broadly used also, as they can be easily implemented with no complex hardware. Another important advantage of this kind of controllers is its fast dynamic response and their inherent capability to limit the peak current injected by the converter. In addition, hysteresis controllers do not require any information about the system parameters, something that enhances its robustness. In classical hysteresis controllers the error band is normally fixed to a certain value. As a consequence the switching frequency varies within a band because peak-to-peak current ripple must be controlled at all points of the fundamental frequency wave. In order to solve this problem without damaging the good dynamical performance of the hysteresis current control, an adaptive hysteresis band able to maintain a constant switching frequency is proposed in this paper. In the following the algorithm that has been developed in order to calculate the value of the hysteresis band will be detailed. Later on their performance will be tested in three different power converter topologies used for single phase PV applications. The results obtained

with the new algorithm will be compared with those obtained when using classical hysteresis and PR controllers. This analysis will be extended as well to the thermal losses that appear in the power converter in each case.

II. BASIC HYSTERESIS CURRENT CONTROL

The basic hysteresis current control is based on an on-line PWM control that fixes the output voltage of the inverter instantaneously [5]. The main task of the PWM current controller in an inverter is to adjust the output current, i , in order to track the current reference provided by i^* . Comparing the instantaneous current in the load with the reference signal the controller should adjust the duty cycle of the PWM signal in the inverter. As a consequence, the error signal (δ) should be reduced [6]. A basic scheme of the PWM current controller is shown in Fig. 1.

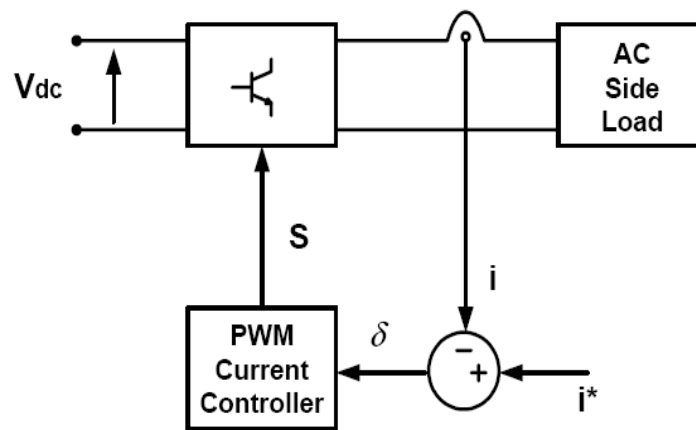


Fig. 1. Basic Current Control Scheme in a single phase inverter.

In this kind of controller, the output voltage level depends on the error between the current setpoint and the real currents injected by the converter. In this way when the load current is lower than the current reference, the inverter connects the positive side of the DC bus to the load, reducing thus the currents. On the contrary when the load current is higher than the current reference, the inverter connects the negative side of the DC bus to the load. Taking into account the previous description, the error signal can be maintained within a certain fixed band. Fig. 2 shows the evolution of the load current when using a basic hysteresis current control.

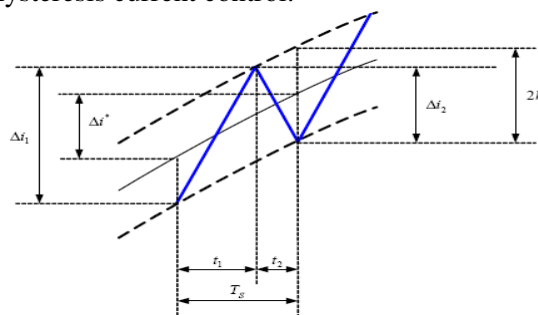


Fig. 2. PV inverter Output current with hysteresis control.

III. PROPOSED HYSTERESIS CURRENT CONTROL

Considering that the output current of the inverter is the one shown in Fig. 2, the error signal can be written as (1):

$$\delta = i - i^* \quad (1)$$

where i is the output current and i^* is the reference current. It is possible to deduct an equation for each switching period written in (2) and (3). According to Fig 2, for $[0, t_1]$:

$$\frac{d\delta}{dt} = \frac{di}{dt} - \frac{di^*}{dt} = \frac{\Delta i_1}{t_1} - \frac{\Delta i^*}{T_s} = \frac{2h}{t_1} \quad (2)$$

And for $[t_2, T_s]$:

$$\frac{d\delta}{dt} = -\frac{di}{dt} - \frac{di^*}{dt} = -\frac{\Delta i_2}{t_2} - \frac{\Delta i^*}{T_s} = -\frac{2h}{t_2} \quad (3)$$

Fig. 3 shows a PV inverter with H5 structure [8]. During the positive semi cycle when switches S5, S1 and S4 are ON and S2, S3 are OFF, an active vector is applied to the load. This switching semi period correspond to t_1 in Fig. 2 when the current is increasing. In order to apply a zero voltage vector to the load, S5 is turn off whereas, S1, S2, S3 and S4 remain in the same state. This switching period correspond to t_2 in Fig. 2 when the current is decreasing (freewheeling case). The equivalent circuits in both periods are shown in Fig 4.

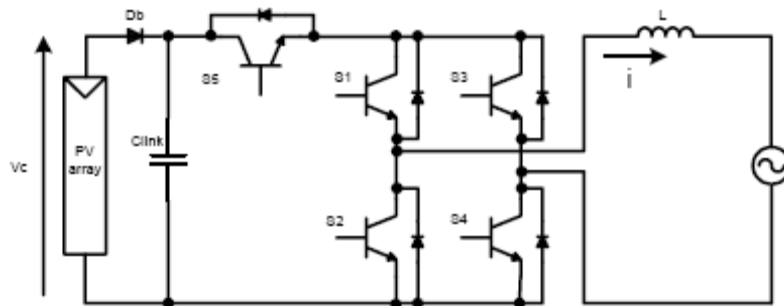


Fig. 3. Basic structure of a H5 PV inverter.

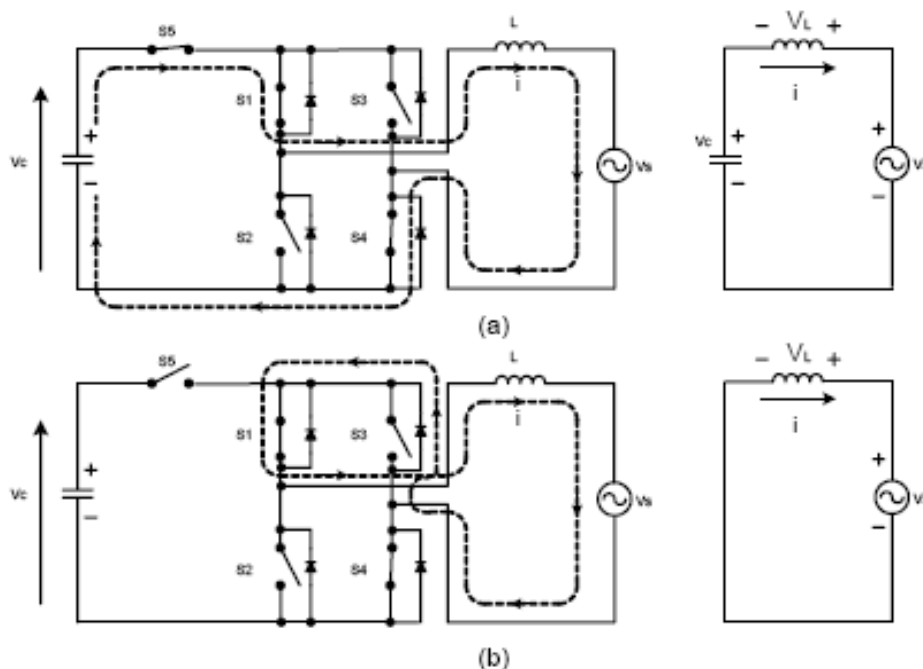


Fig. 4. Equivalent circuits (a) active voltage vector, (b) zero voltage vector.

According with the equivalent circuits shown in Fig. 4, it is possible to establish the equations for the equivalent circuits in both periods (4) and (5).

$$V_C - V_s = L \cdot \left(\frac{di^*}{dt} + \frac{d\delta}{dt} \right) \quad (4)$$

$$0 - V_s = L \cdot \left(\frac{di^*}{dt} + \frac{d\delta}{dt} \right) \quad (5)$$

Substituting equations (2) and (3) in equations (4) and (5) respectively it is possible to obtain the following equations:

$$\left(V_C - V_s - L \frac{di^*}{dt} \right) \cdot t_1 = 2 \cdot L \cdot h \quad (6)$$

$$\left(-V_s - L \frac{di^*}{dt} \right) \cdot t_2 = -2 \cdot L \cdot h \quad (7)$$

Adding equations (6) and (7):

$$t_1 = \frac{T_s}{V_C} \left(V_s + L \frac{di^*}{dt} \right) \quad (8)$$

Finally, by means of substituting the equation (8) in equation(6) and simplifying terms equation (9) is reached:

$$h = \frac{T_s}{2L} \cdot \left(V_s + L \frac{di^*}{dt} \right) \left[1 - \frac{1}{V_C} \cdot \left(V_s + L \frac{di^*}{dt} \right) \right] \quad (9)$$

Equation (9) defines an adaptive hysteresis band that depends on the system parameters. This algorithm permits to obtain a constant switching frequency in the PV inverter load. The same analysis can be performed for other topologies like HERIC (High Efficiency Reliable Inverter Concept) [9], shown in Fig. 5, or HB-ZVR (Half-Bridge Zero Voltage Rectifier) [10], depicted in Fig. 6.

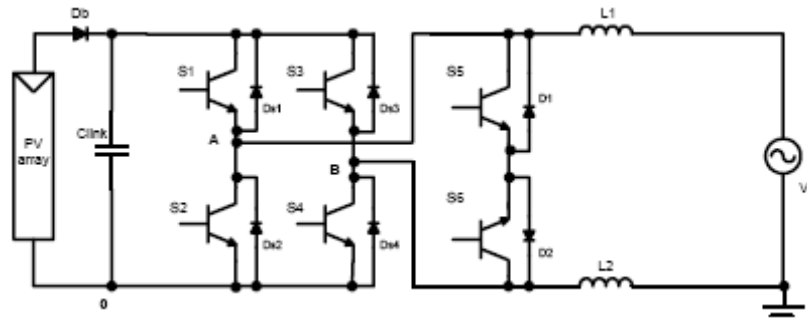


Fig. 5. High Efficient and Reliable Inverter Concept (HERIC) topology.

In the HERIC topology during the positive half cycle S6 remains connected, whereas S1 and S4 commute at switching frequency in order to generate both active and zero vectors. When an active vector is present (S1 and S4 are ON), current flows from the PV panels to the load (grid), when a zero vector occurs, S1 and S4 are switched OFF and then, the current flows through S6 and D1, this is the freewheeling situation. On the other hand, when the negative cycle is coming, S6 goes OFF and S5 goes ON, whereas S3 and S2 commute at switching frequency. This means that an active vector is present when S3 and S2 are ON, therefore the current flows from the PV panel towards the load, thus when S3 and S2 turn off, a zero voltage vector is present in the load, then current flows through S5 and D2. This kind of modulation permits having a freewheeling situation with no connected load (utility grid).

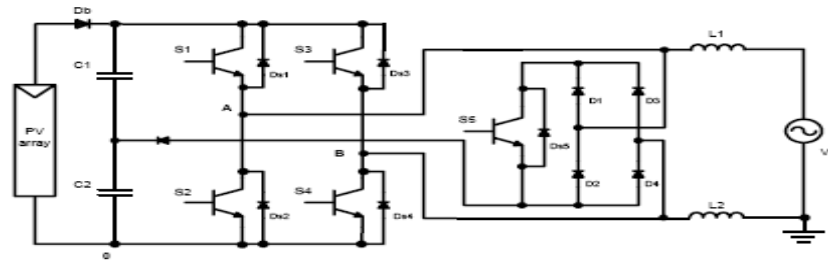


Fig. 6. Half-Bridge Zero Voltage Rectifier (HB-ZVR) topology.

In turn the HB-ZVR topology, Fig. 6, it provides the active vectors using a conventional full-bridge inverter, whereas the zero voltage vectors are generated by the single phase rectifier and S5. This topology works as follow: during positive semicycle, when an active vector is applied, switches S1 and S4 are ON, and S2, S3 and S5 are OFF, in this situation the output current is increasing. In order to apply a zero voltage vector, S1 and S4 are turn off and S5 is turn on, in this case, the current is decreasing. On the other hand, in the negative semi-cycle, an active vector is applied when S2 and S3 are ON, and S1, S4 and S5 are OFF, whereas a zero voltage vector is applied when S1, S2, S3 and S4 are OFF and S5 is ON. In the next section, simulations results using hysteresis current control were performed in order to compare the system performance with and without adaptive hysteresis band. All simulations were performed using PSIM (Thermal Module). The IGBT is an IPM (Intelligent Power Module) from Mitsubishi PM75DSA120. In this study case the DC bus was of 400V, while the load was a grid with $L = 4 \text{ mH}$, $V_s = 325 \text{ Vpk}$ and $f = 50 \text{ Hz}$.

IV. SIMULATION RESULTS

In this section simulation with H5, HERIC and HB-ZVR topologies have been carried out. The simulation results using the adaptive band hysteresis algorithm in the H5, HERIC and ZB-ZVR topologies are shown in Fig. 7, Fig. 8 and Fig. 9 respectively.

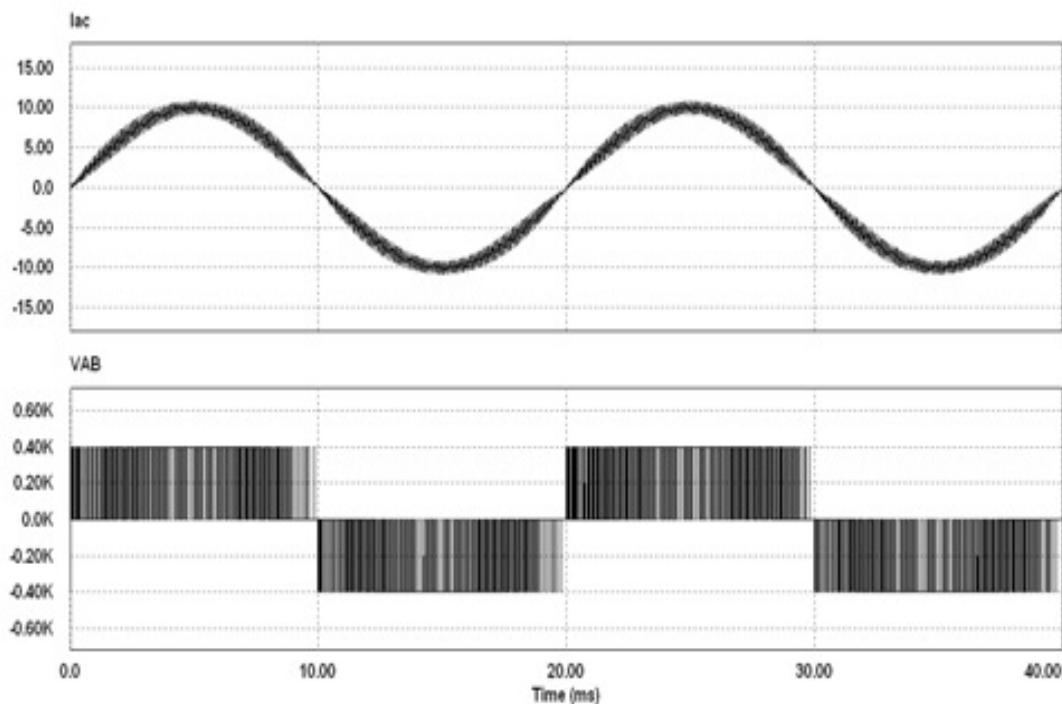


Fig. 7. Behavior of the current and the voltage at the output of the converter when using the H5 topology.

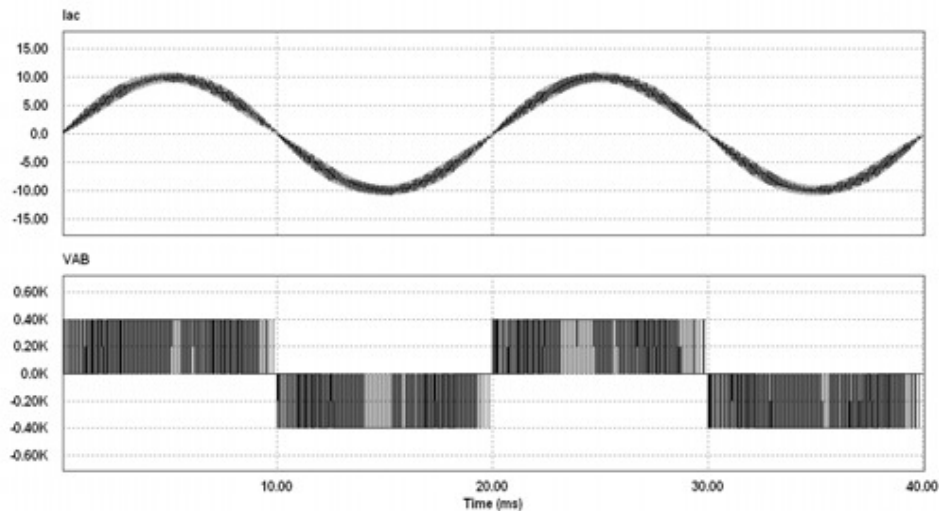


Fig. 8. Behavior of the current and the voltage at the output of the converter when using the HERIC topology.

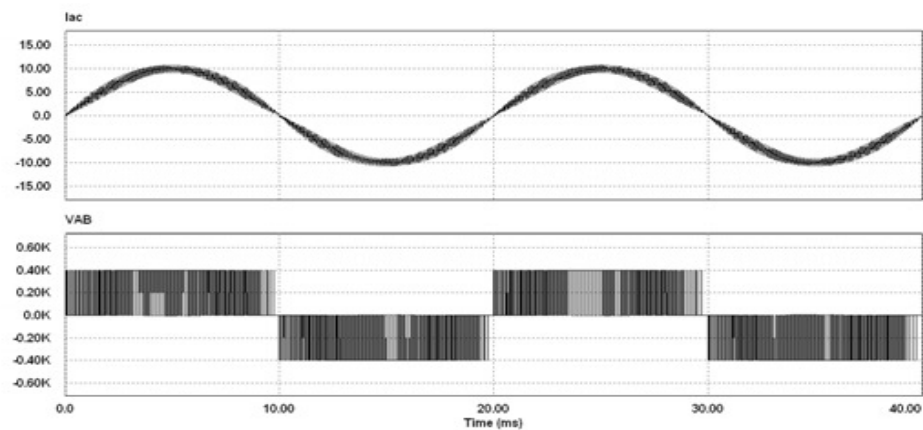


Fig. 9. Behavior of the current and the voltage at the output of the converter when using the or HB-ZVR topology.

In order to see that the switching frequency is almost constant in all three cases, it is necessary to obtain the output current harmonic spectrum. Fig. 10 shows the harmonic spectrum for the HERIC topology with and without adaptive hysteresis band. In the other two cases (H5 and HB-ZVR), the current spectrum is almost the same.

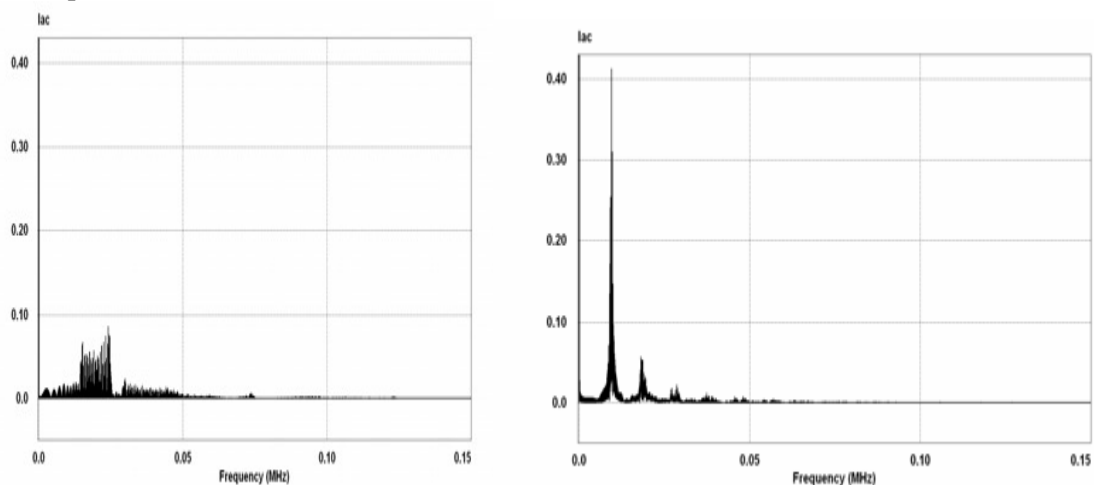


Fig. 10. Frequency spectrum of the output current for HERIC topology, (a) Without adaptive hysteresis, (b) With adaptive hysteresis.

As can be seen in Fig. 10, in case (a) the switching frequency is variable over a wide frequency range and in the case (b) the switching frequency remains around 10 kHz. This characteristic permits to design easily the output filter of the inverter. In the next section it is presented some simulations results regarding switching and conduction losses for these three topologies. The IGBT model used in these simulations was obtained by means IGBT datasheet and after that, was loaded to PSIM library.

V. COMPARATIVE LOSSES ANALYSIS

In order to obtain good comparative losses analysis, all three topologies were simulated with a fixed hysteresis band current control (Basic H), adaptive hysteresis band current control (Adaptive H) and proportional resonant current control (PR Ctrl). Table 1 shows the simulation results of total switching and conduction losses with an input power of 1650 W ($I_{out}=10\text{Amp}$) for all three topologies.

TABLE 1
SWITCHING AND CONDUCTION LOSSES

Topology	$P_{in}=1650W / I_{out}=10Amp$		Control
	Switching Losses (W)	Conduction Losses (W)	
HERIC	47.40	13.74	Basic H
	23.54	12.76	Adaptive H
	25	12.81	PR Ctrl.
H5	26.16	18.08	Basic H
	12.01	18.85	Adaptive H
	13.09	18.05	PR Ctrl.
HB-ZVR	89.02	15.16	Basic H
	46.69	15.1	Adaptive H
	49.53	15.16	PR Ctrl.

As can be seen in Table 1, the major losses come from IGBT switching processes. This means that the switching frequency has an important influence in the total losses of the system. Due this, when the control is implemented with a fixed hysteresis band, the losses are major than when an adaptive hysteresis band current control system is used. As shown in Fig. 10 (a), different frequency harmonics appear over a grid period, this harmonics produce an increase of switching losses. Table 1 shows also, that conduction losses are almost the same in each case. In Table 2 are summarized the distributed losses among all semiconductors in each topology. It is possible to distinguish that the major losses are located in those semiconductors which are switching at high frequency. In addition it is possible to see the specific losses for each semiconductor; with this information, it is possible to perform a good thermal design.

TABLE 2
DISTRIBUTION LOSSES FOR ALL THREE TOPOLOGIES

Topology	Semiconductor/ $P_{in}=1650W / I_{out}=10Amp$											Total Losses	Hysteresis
	S1	S2	S3	S4	S5	S6	D1	D2	D3	D4	D5		
HERIC	13.20	13.26	13.20	13.26	1.09	1.09	2.99	2.99	***	***	***	61.14	Basic
	7.49	7.51	7.49	7.51	1.12	1.12	2.01	2.01	***	***	***	36.31	Adaptive
	7.83	7.85	7.83	7.85	1.12	1.12	2.09	2.09	***	***	***	37.82	PR Ctrl.
H5	3.45	3.45	3.45	3.45	15.21	15.22	***	***	***	***	***	44.25	Basic
	3.11	3.12	3.1	3.1	9.21	9.21	***	***	***	***	***	30.86	Adaptive
	3.21	3.17	3.28	3.3	9.09	9.09	***	***	***	***	***	31.15	PR Ctrl.
HB-ZVR	12.35	12.38	12.36	12.38	42.80	***	2.97	2.97	2.97	2.97	0	104.19	Basic
	7.51	7.53	7.52	7.53	23.68	***	2.01	2.01	2.01	2.01	0	61.86	Adaptive
	7.83	7.84	7.83	7.85	24.98	***	2.08	2.08	2.08	2.08	0	64.7	PR Ctrl.

In the case of H5 topology, the losses are located in S5. In all three cases the losses behavior is almost the same, the losses are reduced when adaptive hysteresis band current control and proportional resonant current control are used.

VI. CONCLUSIONS

A hysteresis current control algorithm based on an adaptive hysteresis band for single phase PV converter topologies has been presented in this paper. As it has been shown analytically and by means of simulations this algorithm permits obtaining a fixed switching frequency in all the tested topologies. The main drawback of the conventional fixed hysteresis band current control is that generates excessive current ripple because modulation frequency varies within a band. This

modulation frequency variation makes complicated the output filter design. Adaptive hysteresis band current control keeps the good performance of the fixed band hysteresis current control and additionally permits an easier output filter design due that the switching frequency is almost constant. On the other hand, switching losses can be reduced by using this adaptive hysteresis band current control. The analyzed topologies are the more widely used in transformerless single-phase PV systems (H5 and HERIC). Based in the previously comparative simulations results it can be concluded that in the case of H5 topology losses are concentrated in S5. In case of HERIC topology losses are located among S1, S2, S3 and S4. Finally in HB-ZVR single phase topology, losses are located in S5. These results mean that in each case, the losses distribution is not the same and a different thermal design should be done.

REFERENCES

- [1] L. Malesani, P. Mattavelli, P. Tomasini, "High Performance Hysteresis Modulation Technique for Active Filters", *IEEE Transactions on Power Electronics*, Volume 12, September 1997.
- [2] J. Holtz and S. Stadtfeld, "A Predictive Controller for the Stator Current Vector of AC Machine-fed from a Switched Voltage Source", in Proc.Int. Power Electronics Conference Rec. (Tokyo), 1983, pp. 1665-1675.
- [3] M. Ciobotaru, R. Teodorescu, and F. Blaabjerg, "Control of Single-Stage Single-Phase PV Inverter", *European Conference on Power Electronics and Applications*, 2005.
- [4] Y. Hayashi, N. Sato, K. Takahashi, "A Novel Control of a Current- Source Active Filter for ac Power System Harmonic Compensation", *IEEE Transactions on Industrial Applications*, Vol. 27, No. 2, March/April 1991.
- [5] T. Kato, K. Miyao, "Modified Hysteresis Control with Minor Loops for Single-Phase Full-Bridge Inverters", Doshisha University, Kyoto Japan, 88CH2565-0/88/0000-0689\$01.00, 1988 IEEE.
- [6] Bimal K. Bose, "An Adaptive Hysteresis- Band Current Control Technique of a Voltage-Fed PWM Inverter for Machine Drive System", *IEEE Transactions on Industrial Electronics*, Vol. 37, No 5, October 1990.
- [7] Marian P. Kazmierkowski, Luigi Malesani, "Current Control Techniques for Three-Phase Voltage-Source Converters: A Survey", *IEEE Transactions on Industrial Electronics*, Vol. 45, No. 5, October 1998.
- [8] EP 1 626 494 A2; Europäische Patentanmeldung, European Patent Office
- [9] EP 1 369 985 A2; Europäische Patentanmeldung, European Patent Office
- [10] T. Kerekes, G. Vazquez, P. Rodriguez, E. Aldabas, R. Teodorescu, "A New High-Efficiency Single-Phase Transformerless PV Inverter Topology", *IEEE Transactions on Industrial Electronics*, 2009

Authors Information

B. NAGARAJU received his M.Tech (Power Electronics, 2009) from, Vaagdevi College of Engineering, Warangal. He is currently working as an Assistant Professor in Department of EEE; Vaagdevi College of Engineering. His areas of interest include Power Quality Maintenance in Smart Grid using Renewable Energy Sources.



K. PRAKASH received his M.Tech (Power Systems, 2003) from National Institute of technology, Warangal. He is currently pursuing his Ph.D (Electrical Engineering) National Institute of technology, Warangal. His areas of interest include Distribution System Studies, Economic Operation of Power Systems, Artificial Intelligence Techniques and Meta-Heuristics Techniques.



PARAMETRIC STUDY OF A NOVEL STACKED PATCH ANTENNA

V. Rajya Lakshmi¹, M. Sravani², G.S.N.Raju³

¹Department of Electronics and Communications Engineering,
Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam

^{2&3}Department of Electronics and Communications Engineering,
Andhra University College of Engineering (A), Visakhapatnam

ABSTRACT

An inverted stacked microstrip patch antenna with single-fed at 1.30 GHz for GPS applications is designed at L band using electromagnetic simulator WIPL-D. The radiation patterns of microstrip antennas depend upon parameters like width, length, resonant frequency and dielectric constant. In view of this, by varying the height of the air gap, feed position and dielectric constant of the patch, corresponding radiation pattern, near field and return loss are observed.

KEYWORDS: Stacked microstrip patch antenna, Return loss, Radiation pattern, WIPL-D

I. INTRODUCTION

Microstrip antennas often called patch antennas [1-5] are widely used in the microwave frequency region [6-7]. The advantages of the microstrip antennas are small size, low profile, and lightweight, conformable to planar and non planar surfaces. It demands a very little volume of the structure when mounting. They are simple and cheap to manufacture using modern printed circuit technology. There are many configurations that can be used to feed microstrip antennas. The four most popular are microstrip line, coaxial probe, aperture coupling and proximity coupling [8].

A wide range of substrate materials is available, clad with copper, aluminum or gold. Low cladding thickness simplify fabrication of the antenna to required tolerances, whereas thicker clads ease soldering [9]. For high power applications of microstrip antennas, a thick cladding is desirable. The large range of substrates are available, viz., PTFE, Polystyrene, Polyphenylene, Ceramic, etc., permits considerable flexibility in the choice of a substrate for particular applications [10].

In this paper, an inverted stacked microstrip patch antenna is designed with microstrip line feed using WIPL-D at frequency 1.30 GHz for GPS applications. GPS is global positioning system based on satellite technology [11-19]. In microstrip line type of feed technique, a conducting strip is connected directly to the edge of the microstrip patch. The conducting strip is smaller in width as compared to the patch and this kind of feed arrangement has the advantage that the feed can be etched on the same substrate to provide a planar structure. Hence this is an easy feeding scheme, as it provides ease of fabrication and simplicity in modeling. In order to simplify analysis and performance prediction, the patch is generally square, rectangular, circular, triangular, and elliptical or some other common shape. The effect of various parameters such as relative permittivity (ϵ_r), the distance of feed position from center, and height of the air gap (h) on the performance of the antenna has been studied.

II. STACKED MICROSTRIP PATCH ANTENNAS

Coupled patch antennas can be made of two or multiple patches. The case of two stacked coupled patches can be viewed as a single wideband microstrip patch antenna that can be used alone or in an array [20-22]. Its gain will be slightly larger than that of a single-layered patch, due to its increased height above the ground plane, and controlled by its value in wavelength; that is, the larger the height of the stacked configuration, the larger its gain. This is the secondary benefit from the stacked structure, in addition to its broadband impedance bandwidth. The case of multiple coupled patch antennas can increase significantly both the input impedance bandwidth and the gain.

The coupled patch antennas have fundamentally two different configurations, the planar structures and stacked configurations. Both these configurations were investigated concurrently. In the planar configuration the coupling is from the edges of the adjacent patches. Since the geometry represents two coupled resonators, the coupling coefficient is the most important parameter for broad banding of the input impedance [23-24]. However, the coupling in this case is normally weak and below the coupling level necessary for optimum broad banding. Thus the degree of bandwidth broadening is limited to around 10 percent, unless multiple planar patches are used. Nevertheless, this type of impedance bandwidth increase becomes attractive in the design of planar arrays, like microstrip Yagi arrays, where the planar structure makes this type of impedance broadening very natural. In other cases this configuration becomes appealing only if there is a limitation on the antenna height and adequate planar surface is available for the coupled patches.

The stacked coupled patch structures have been favored in most applications because of the ease with which the coupling coefficient can be controlled by adjusting their separation height. Their other parameters such as the size of the patches and substrate parameters can also be modified independently to facilitate the designs [25-26]. Consequently, much research has been done on these antennas, and impedance bandwidths of 10–30 percent for (voltage standing wave ratio) VSWR < 2 have been achieved. From broadband network theory, if the coupling between the two patches is too strong, i.e., above the required critical coupling, there are two resonance frequencies that increasingly separate by increasing the coupling coefficient. Since in stack patches their spacing controls the coupling, the antenna becomes dual band when patches are too close to each other. As the spacing is increased, the coupling reduces and the two resonances merge at the critical coupling to provide a broadband antenna. Further increase of the spacing reduces the bandwidth gradually.

III. ANTENNA DESIGN

The basic design considerations for a rectangular patch were considered for the design. The length L of the patch is usually $0.3333\lambda_0 < L < 0.5\lambda_0$, where λ_0 is the free-space wavelength. The patch is selected to be very thin such that $t \ll \lambda_0$ (where t is the patch thickness). The height h of the dielectric substrate is usually $0.003\lambda_0 \leq h \leq 0.05\lambda_0$. The dielectric constant of the substrate (ϵ_r) is typically in the range $2.2 \leq \epsilon_r \leq 12$. Formulas used for the patch design is given below [27-32].

$$\text{Patch width} = w = \frac{c}{2f \sqrt{\frac{(\epsilon_r + 1)}{2}}} \quad (1)$$

$$\text{Patch length} = L = L_{eff} - 2dL \quad (2)$$

$$\text{Where } L_{eff} = \frac{c}{2f \sqrt{\epsilon_{reff}}} \quad (3)$$

$$dL = 0.412h \frac{(\epsilon_{reff} + 0.3) \left(\frac{W}{h} + 0.264\right)}{(\epsilon_{reff} - 0.258) \left(\frac{W}{h} + 0.8\right)} \quad (4)$$

$$\epsilon_{reff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left(1 + 12 \frac{h}{W}\right)^{-\frac{1}{2}} \quad (5)$$

$$\text{Ground length, } L_g = 6h + L \quad (6)$$

$$\text{Ground Width, } W_g = 6h + W \quad (7)$$

The proposed antenna structure is shown in Fig 1. The top patch (length p_2) is printed on substrate with thickness t_2 and dielectric constant ϵ_{r2} , and the bottom patch (length p_1) printed on substrate with thickness t_1 and dielectric constant ϵ_{r1} , while a air gap with thickness h separate the two layers of substrates, L is length of the ground plane. Polyflon–Norclad ($\epsilon_{r2} = 2.55$, $\tan\delta = 0.0011$) with $t_2 = 1.5\text{mm}$ as superstrate [33]. FR4-epoxy ($\epsilon_{r1} = 4.4$, $\tan\delta = 0.02$) with $t_1 = 1\text{mm}$ is used as substrate. In this paper, corner truncated approach is used with a pair of opposite corners (side length q_1 , q_2). In stacked patch antenna, h is the other freedom parameter affecting return loss performance, which controls the coupling between patches [34–35].

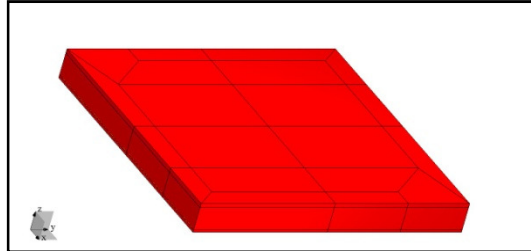


Fig 1: Geometry of the proposed stacked microstrip antenna

Fig 2: shows the cross sectional view of the microstrip patch antenna designed in WIPL-D with line feeding and table 1: Shows the designed parameters of patch antenna using line feeding.

Table 1: Parameters of the patch antenna using line feeding (mm)

L	p1	p2	q1	q2	t1	t2	h
97.6	52.6	82.6	7.2	14.8	1	1.5	8

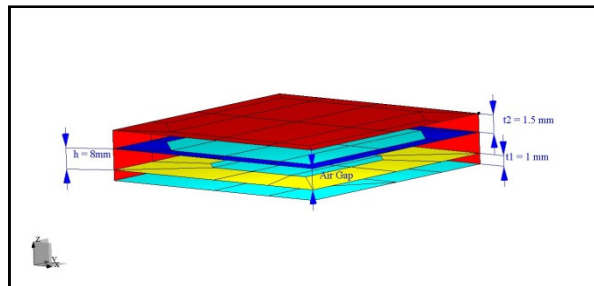


Fig 2: Cross sectional view of microstrip patch antenna

Figs 3 and 4 indicate the upper (parasitic) and lower (radiating) patches of proposed stacked patch antenna.

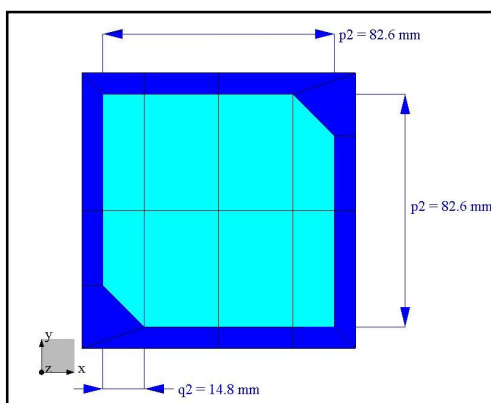


Fig 3: Parasitic patch

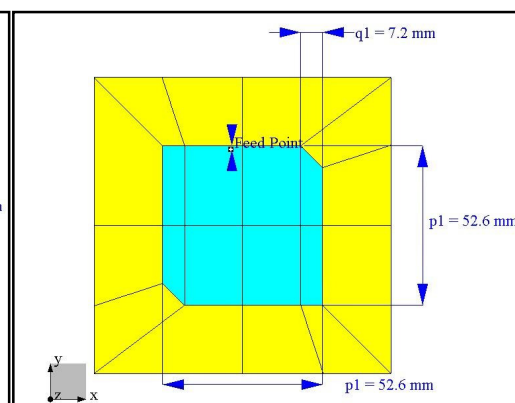


Fig 4: Radiating patch

Antenna consists of ground plane, square patch and dielectric parts. Cyan color indicates the metallic surfaces, Yellow, dark blue and Red indicate dielectric surfaces of first and second layers respectively [36].

IV. MODELING AND SIMULATION

The software used to model and simulate the microstrip patch antenna is WIPL-D. It is a 3D frequency domain EM solver based on method of moments [37-38]. It enables to model the structures of arbitrary shape using wires and plates as basic building blocks. The basic steps when using the 3D EM solver are defining the problem, editing and checking the input data, saving the project and running the analysis, listing and plotting the computed results [39-40].

V. RESULTS

Fig. 5 represents the overlaid Return Loss at different dielectric constants, maximum Return Loss observed at $\epsilon_{r2}=2.55$ Fig. 6 shows the overlaid Return Loss at corresponding heights, maximum Return Loss observed at $h=8\text{mm}$. Fig. 7 shows the overlaid Return Loss of corresponding feed positions, maximum Return Loss observed at feed= $(-4,25)$. Figs. 8-10 represent the overlaid Radiation pattern at $\phi=180^\circ$ at different dielectric constants, heights and feed positions respectively. Fig. 11 shows the Radiation pattern in 3D at optimum feed = $(-4, 25)$, $h=8\text{mm}$, $\epsilon_{r2}=2.55$. Tables 2-4 represent the Return Loss observed at different values of ϵ_{r2} , feed position and height respectively.

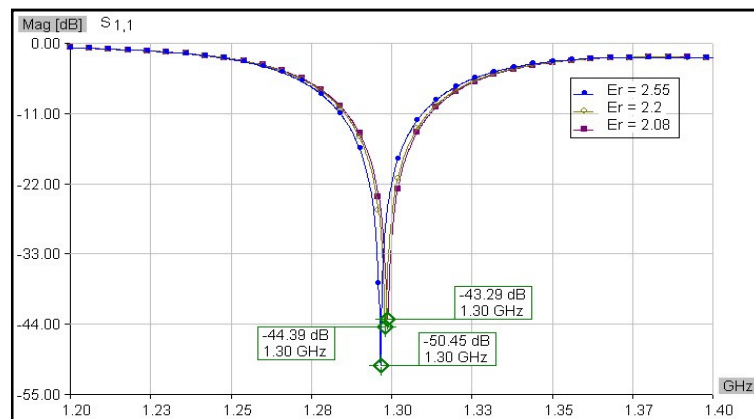


Fig 5: Return Loss at $\epsilon_{r2} = 2.55, 2.2, 2.08$

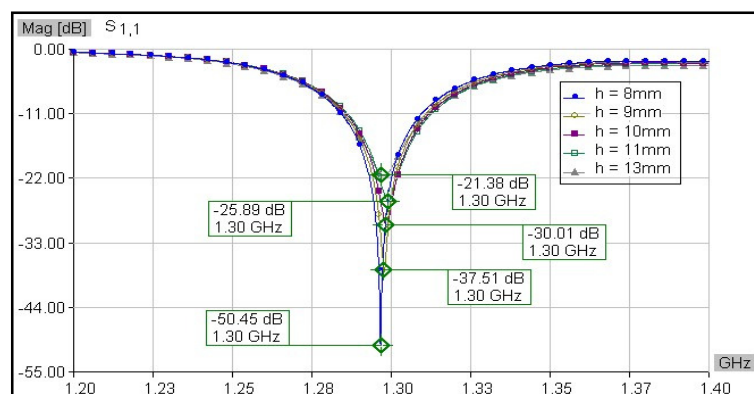


Fig 6: Return Loss at $h=8\text{mm}, 9\text{mm}, 10\text{mm}, 11\text{mm}, 13\text{mm}$.

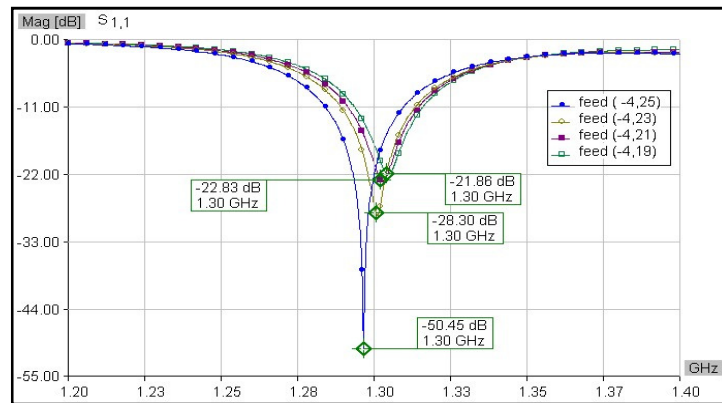


Fig 7: Return Loss at feed position = (-4, 25), (-4, 23), (-4, 21), (-4, 19)

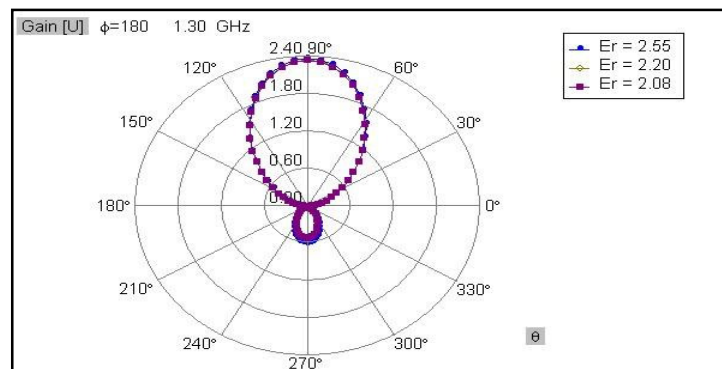


Fig 8: Radiation pattern at $\epsilon_{r2} = 2.55, 2.2, 2.08$

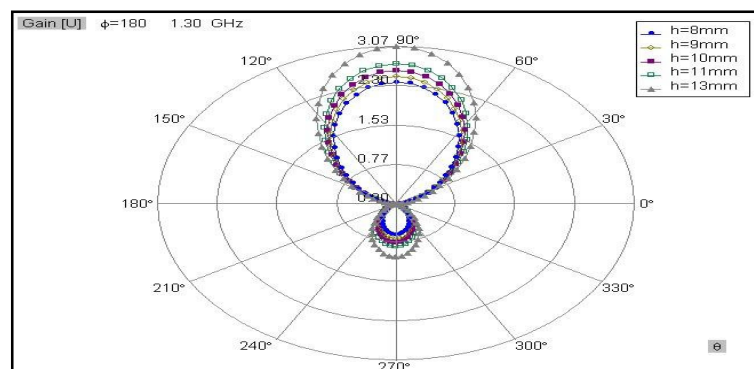


Fig 9: Radiation pattern, at $h=8\text{mm}, 9\text{mm}, 10\text{mm}, 11\text{mm}, 13\text{mm}$

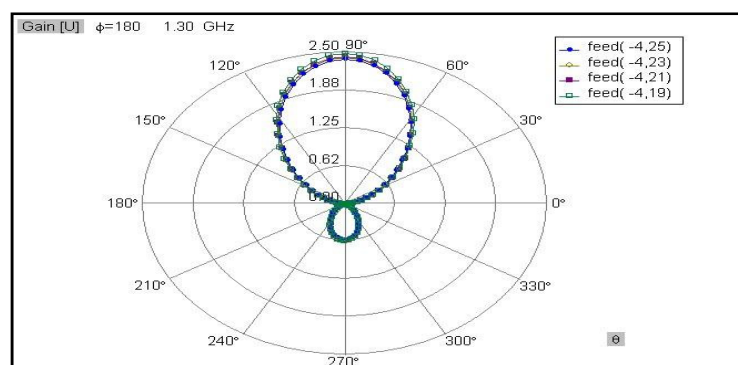


Fig 10: Radiation pattern, at feed position = (-4, 25), (-4, 23), (-4, 21), (-4, 19)

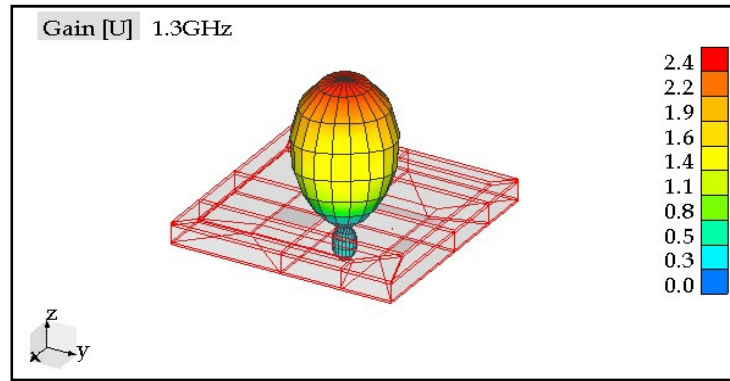


Fig 11: Radiation pattern observed in 3D at feed = (-4, 25), h=8mm, $\epsilon_{r2}=2.55$

Table: 2 Return Loss at different values of ϵ_{r2}
 $\epsilon_{r1}=4.4$, h=8mm, feed position (-4, 25)

ϵ_{r2}	Return Loss (dB)	Frequency (GHz)
11.9	-19.07	1.30
4.882	-32.38	1.30
9.5	-21.48	1.30
2.08	-43.29	1.30
2.2	-44.39	1.30
2.55	-50.45	1.30
6	-27.53	1.30
10.8	-19.99	1.30

Table: 3 Return Loss observed at different values of feed position
At $\epsilon_{r1}=4.4$, $\epsilon_{r2}=2.55$, h=8

Feed position	Return Loss in dB	Frequency (GHz)
-4,19	-21.82	1.30
-4,21	-22.83	1.30
-4,23	-28.30	1.30
-4,25	-50.45	1.30

Table: 4 Return Loss observed at different values of height
At $\epsilon_{r1}=4.4$, $\epsilon_{r2}=2.55$, feed position (-4, 25)

Height (mm)	S11(RL) in dB	Frequency (GHz)
7	-33.26	1.30
8	-50.45	1.30
9	-37.51	1.30
10	-30.03	1.30
11	-25.88	1.30
12	-24.01	1.30
13	-21.38	1.30
14	-19.37	1.30

VI. CONCLUSION

An inverted patch antenna is designed with double layers of truncated square patches using line feeding to operate at L-band for GPS applications using WIPL-D. By varying the feed position, dielectric constant and height of the air gap (h) corresponding radiation pattern and return loss are observed. The optimum feed position, dielectric constant and height of the air gap (h) are found to be

(-4, 25), $\epsilon_{r2} = 2.55$ and $h = 8\text{mm}$ respectively. The return loss of -50.45 dB is observed at designed frequency 1.30GHz.

REFERENCES

- [1] G.S.N. Raju, Antennas and Wave Propagation, Pearson Education, 2006.
- [2] Pozar, D.M. and Schaubert, D.H., Microstrip Antennas, IEEE Press, Piscataway, NJ, 1995.
- [3] Carver, K. R., Mink, J. W. Microstrip Antenna Technology, IEEE Transactions on Antennas and Propagation, v. AP-29, Jan. 1981, pp. 2-24.
- [4] R. Waterhouse, Small microstrip patch antenna, Electron. Lett. 31, 604–605, April 13, 1995.
- [5] S. Dey and R. Mittra, Compact microstrip patch antenna, Microwave Opt. Technol. Lett.13, 12–14, Sept. 1996.
- [6] Shancheng Pan, and Kinlu Wong, Dual-Frequency Triangular Microstrip Antenna with a Shorting Pin, IEEE Transactions on Antennas and Propagation, vol. 45, pp.1889-1891, 1997.
- [7] Jaume Anguera, Carles Puents, and Carmen Borja, Dual-frequency Broad-Band stacked Microstrip patch Antennas, IEEE. Antennas Wireless Propag. Lett , vol. 2, pp. 36-39, 2003.
- [8] Constantain. A. Balanis, Antenna Theory Analysis and Design, III Edn., Wiley-Inter science. A John Wiley & Sons Inc. Publication.
- [9] R. B. Waterhouse, Stacked Patches using High and Low Dielectric Constant Material Combinations, IEEE Transactions on Antennas and Propagation, Vol. 47, pp. 1767-1771, Dec. 1999.
- [10] J. Bahl and P. Bhatia, Microstrip Antennas, Artech House, Norwood, MA, 1980.
- [11] Global View, GPS World Magazine, February 2002, p. 10.
- [12] Understanding GPS: principles and applications By Elliott D. Kaplan, Christopher J. Hegarty.
- [13] Introduction to GPS: the Global Positioning System, Ahmed El-Rabbany.
- [14] Global Positioning System: theory and practice, B. Hofmann-Wellenhof, Herbert Lichtenegger, James Collins.
- [15] The global positioning system: theory and applications, Volume 2; Volume 164 James J. Spilker
- [16] Guide to GPS positioning, David Wells, Norman Beck, Canadian GPS Associates
- [17] Global Positioning System: Principles And Applications, Satheesh Gopi
- [18] Zhizhang Chen, Acep D. Ganjara, and Xiaomin Chen, A Dual-L Antenna with a novel tuning technique for dual frequency applications, IEEE Transactions on Antennas and Propagation, vol. 50, pp. 402-403, March 2002.
- [19] Guochang Xu, GPS: Theory, Algorithms and Applications, Second edn., Springer.
- [20] Long S. A., Walton, M. D., A Dual-Frequency Stacked Microstrip Antenna, IEEE Transactions on Antennas and Propagation, v. AP-27, n. 2, 1979, pp. 270-273.
- [21] Nasimuddin, Karu P. Esselle, and A. K. Verma, Wideband circularly polarized stacked microstrip antennas, IEEE Antennas and wireless Propagation letters, vol. 6, pp. 21- 24, 2007.
- [22] N. Herscovici, Z. Sipus, and D. Bonefaci, Circularly polarized singlefed wideband microstrip patch, IEEE Transactions on Antennas and Propagation, vol. 51, no. 6, pp. 1277-1280, 2003.
- [23] Jianfeng Ye, Yufeng, Wang, and Weizheng Pang, Optimized Design of Circularly Polarized Wideband Stacked Microstrip Antenna With Single- Fed, 2008 International Conference on Microwave and Millimeter Wave Technology (ICMMT 2008), Nanjing vol. 4, pp. 1651-1653, April 2008.
- [24] Kwok Lun Chung, and Ananda Sanagavarapu Mohan, A systematic design method to obtain broadband characteristics for single-fed electromagnetically coupled patch antennas for circular polarization, IEEE Transactions on Antennas and Propagation, vol. 51, pp 3239 -3248, Dec, 2003.
- [25] Hassan Ghannoum, Serge Bories and Christophe Roblin, Probe Fed Stacked Patch Antenna for UWB Sectoral Applications, IEEE Transactions Antennas & Propagation, pp.97-102, 2005.

- [26] R. B. Waterhouse, Broadband stacked shorted patch, Electron. Lett. 35, 98–100, Jan. 21, 1999.
- [27] Pozar, D. M., Microstrip Antennas, Proceedings of the IEEE, v. 80, n.1, Jan. 1992.
- [28] James, J. R. et al. Microstrip Antenna Theory and Design, Peter Peregrinus Ltd, London, 1981.
- [29] James, J.R., Hall, P. S. (eds.) Handbook of Microstrip Antennas, Peter Peregrinus Ltd. (London) on behalf of the IEE, v. 1 & 2, 1989.
- [30] Deschamps, G.A, Microstrip Microwave Antennas, 3rd USAF Symposium on Antennas, 1953.
- [31] K. L. Wong, Design of Nonplanar Microstrip Antennas and Transmission Lines, Wiley, New York, 1999.
- [32] J. D. Kraus, Antennas, McGraw-Hill, New York, 1988.
- [33] Yufeng Wang, Jingbo Cui, Xiaolong Yang, A Dual-Band Circularly Polarized Stacked Microstrip Antenna with Single-fed for GPS Applications, Jiangnan Electronic and Communications Research Institute. IEEE Transactions Antenna Propagation and EM Theory, pp. 108-110, March. 2009.
- [34] R. Chair, K.M. Luk, and K. F. Lee, Small dual patch antenna, Electron. Lett. 35, 762–764, May 13, 1999.
- [35] K. M. Luk, R. Chair, and K. F. Lee, Small rectangular patch antenna, Electron. Lett. 34, 2366–2367, Dec. 10, 1999.
- [36] B. Kolundzija, J. S. Ognjanovic, T. K. Sarkar, WIPL-D: Electromagnetic Modeling of Composite Metallic and Dielectric Structures - Software and User's Manual, Artech House Inc., 2000.
- [37] B. M. Kolundzija, Electromagnetic modeling of composite metallic and dielectric structures, IEEE Trans. Microwave Theory Tech., vol. MTT-47, pp. 1021-1032, July 1999.
- [38] WIPL-D Microwave: Circuit And 3D EM Simulation For RF & Microwave Applications: Software And Users Manual Jovan S. Ognjanovic, Branko Kolundzija, Tapan K Sarkar
- [39] WIPL-D website: <http://wipl-d.com>
- [40] Tabet, S: Antenna Modeling and Analysis Using WIPL-D, Proc. of the 19th Applied Computational Electromagnetics Conf., CD ROM Edition: S17P04.pdf, Monterey, 24-28 March 2003.

Authors

V. Rajya Lakshmi currently pursuing her Ph.D in Andhra University. She is working as Associate Professor in ANITS.



M.Sravani completed her M.Tech in Radar and Microwave Engineering from Andhra University.



G.S.N. Raju received his B.E., M.E. with distinction and first rank from Andhra University and Ph.D. from IIT, Kharagpur. He is Professor of Electronics and Communication Engineering in Andhra University, College of Engineering, Visakhapatnam, India. He is in teaching and research for the last 30 years in Andhra University. At present he holds the positions of Principal, Andhra University College of Engineering (A), Visakhapatnam, Chief Editor of National Journal of Electromagnetic Compatibility.



BIOMETRICS STANDARDS AND FACE IMAGE FORMAT FOR DATA INTERCHANGE - A REVIEW

Nita M. Thakare¹ and V. M. Thakare²

¹Department Computer Science and Engg., S.S.G.M. College of Engg., Shegaon (M.S.), India

²Department of Computer Science and Engg., S.G.B.A. University, Amravati (M.S.), India

ABSTRACT

A biometric system is essentially a pattern recognition system. It compares the feature set data against the set of templates. Initially the biometric data is acquired from the subject, and from the acquired data the feature sets are extracted, these feature sets are used for comparison. With the ever-growing use of biometrics, it is the utmost need to use standard biometric systems. The biometric standards simplifies otherwise complicated choices, enables large scale integration, promote longevity and enables interoperability of the biometric systems. ISO/IEC has developed the biometric standards for all modalities. The part-5 of it contains the Face-Image Format for Data Interchange. It defines specifically a standard scheme for codifying data, describing human faces within a compliant data-structure. In order to enable applications that run on a variety of devices, including those with limited resources and to improve face recognition accuracy; the specification describes not only the data format, but also additional requirements, namely: scene constraints; photographic properties; digital image attributes.

KEYWORDS: Biometric standards, Biometric modalities, Face recognition.

I. INTRODUCTION

Biometric is the technique used to identify or verify a person by using the biological and behavioural characteristics of a person. The main objectives to implement the biometric system are to provide security towards the access of data, prohibit the entry of the unauthorised person in the restricted area, identify a person with criminal record or maintain the working time-record of an employee. It is the automated system which recognizes the individual based on the measurable biological and behavioural characteristics. The characteristics must be any automatically measurable, robust and distinctive physical trait. Biometrics' which are commonly used includes: Fingerprint Recognition, Iris Recognition, Face Recognition, Hand Geometry Recognition, Vein Recognition, Voice Recognition and Signature Recognition. In other words the biometrics can be referred to an automated system that can identify an individual by measuring their physical and behavioral uniqueness or patterns, and comparing it to those on record, instead of requiring personal identification cards, magnetic cards, keys or passwords, biometrics can identify fingerprints, face, iris, palm prints, signature, DNA, or retinas of an individual for easy and convenient verification [1][2]. A biometric system is essentially a pattern recognition system that operates by acquiring biometric data from an individual, extracting a feature set from the acquired data, and comparing this feature set against the template set in the database.

The overall process can be implemented in two phases; Enrolment phase and Authentication phase. At the enrolment end a sample of the biometric trait is captured, processed by a computer, and stored for later comparison. And at the authentication or recognition end biometric system recognizes a person or authenticates a person's claimed identity from their previously enrolled pattern as shown in figure 1. In this process the comparison is carried out by one of two types of searches. They provide

either a one-to-one (A sample is compared with single stored template) or a one- to-many (A sample is searched against database of templates) search capability. One-to-one process which is also known as verification or authentication checks the validity of a claimed identity by comparing a verification template to an enrolment template. One-to-many search is known as identification or recognition, is designed to determine identity of a person based on biometric information only. The template of one person is compared against all biometric templates enrolled in the system.

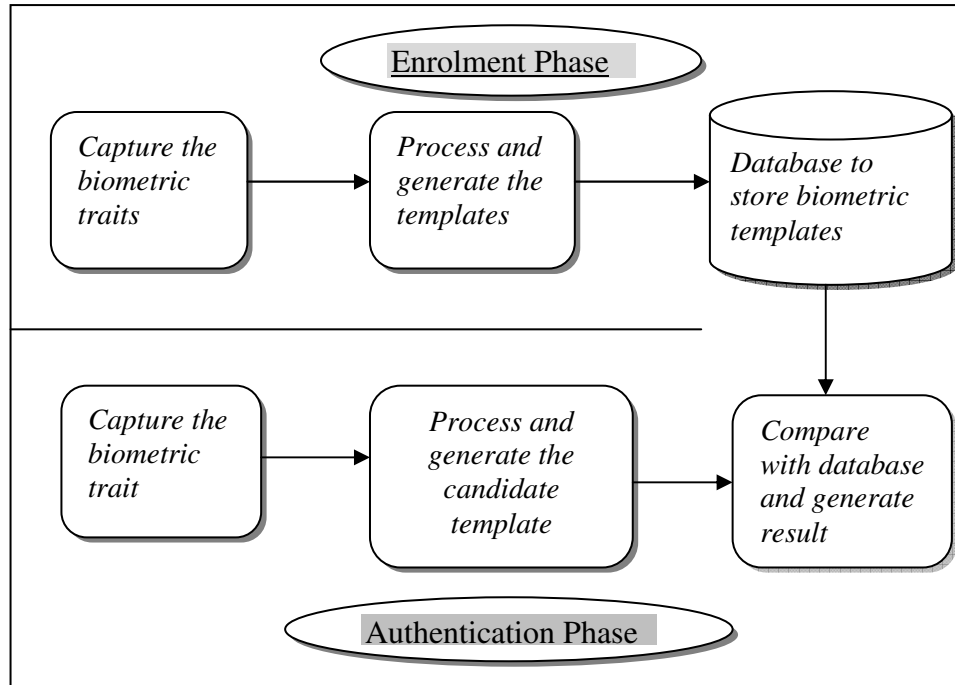


Figure 1: Biometric process

II. THE MEASURES OF MODALITY

The modalities that are being commonly used can be classified into two categories; physiological and behavioural. Physiological biometrics includes fingerprint recognition, iris recognition, retina scan, face recognition, DNA biometrics whereas the voice recognition, signature verification, keystroke verification and gait recognition are considered as behavioural biometrics. For every biometric technique the four possible outcomes are Genuine Accept, False Accept (Error), Genuine Reject, False Reject (Error). The performance of biometric system is evaluated by considering two factors; False Accept Rate (FAR) and False Reject Rate (FRR). The FAR is the chance that someone other than you is granted access to your account. Low false acceptance rate is most important when security is the priority. The FRR is the probability that the genuine user is not authenticated to access his/her account. Low FRR is required when convenience is the important factor. Therefore the balance between these two errors is required to implement an efficient biometric system. To enhance the performance of biometric system as well as to answer the accessibility issues the multimodal biometric is the widely accepted solution. The multimodal biometric system makes use of multiple biometrics like face recognition with signature verification, fingerprint recognition with iris scan, gait recognition with retina scan. The multimodal biometric systems are rapidly progressing. As far as biometrics for personal recognition is concerned, any biological or behavioural characteristic can be used as a biometric identifier providing it satisfy the basic requirements [3]. These requirements are considered as measures of modality.

Universality: Every person should have the characteristic which is considered as the biometric trait for the recognition. The dumb candidate or a person without a fingerprint, need to be accommodated in some other way.

Uniqueness: Generally, no two people have identical characteristics. However, identical twins are hard to distinguish. The combination of physiological and behaviour traits can help to improve identification performance.

Permanence: The characteristics should not vary with time. A person's face, for example, may change with age. In such cases the registration/ enrolment should be repeated after certain period.

Collectability: The characteristics must be easily collectible and measurable.

Performance: The method must deliver accurate results under varied environmental circumstances.

Acceptability: The general public must accept the sample collection routines. Nonintrusive methods are more acceptable.

III. BIOMETRIC STANDARDS

Biometric systems work on images to measure and record biological characteristics. The enrolled/ trained images are converted into templates using image processing algorithms. These templates further are used to verify, recognize or identify individuals. These systems are implemented in mainly four phases; these phases are implemented to collect, store, analyze and exchange data efficiently and securely (see Figure 2). This figure is reconstructed from Biometric Standards for DoD Operational Requirements by [4], it depicts the generalized method of biometric standardization. Biometric standards allow the systems to implement the standardization of various components. It allows ease of maintenance, interoperability, and longevity of systems.

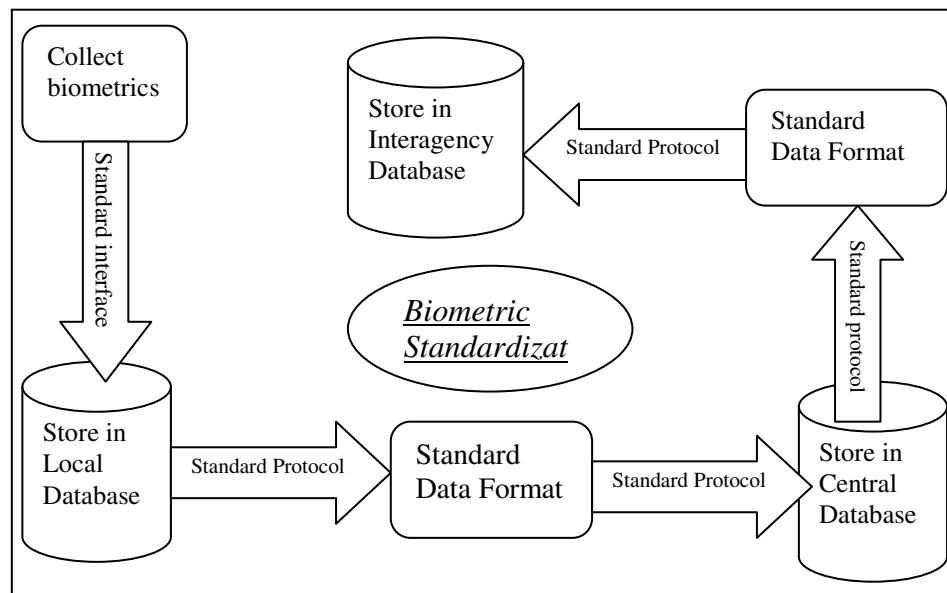


Figure 2. *Biometric Standardization*

International Organization for standardization (ISO)/ International Electrotechnical Commission (IEC) guide 2:2004 defines a standard as “a document, established by consensus that provides rules, guidelines or characteristics for activities or their results” [5]. As shown in the figure-2 biometric standardization uses the standard interfaces, standard protocols and standard data format which makes the biometric system as a standard system. These standards-based interfaces and data formats enable the system to be vendor independent, it incorporates interoperability and data sharing among various subsystems. It also helps to review the existing standards and, if required, modify or extend, to specify the design parameters of the standards and to generate the standard performance results.

The biometric standards support standardized performance and conformance testing which results into uniform test results and predictable matching performance also the standards-based sample quality assessment generates the uniform quality scores, it also help consistent design of quality measurement tools which results into improved matching performance. The four basic types of standards are:

3.1 Interfaces for Technical Information

Technical interface standards are specific to interactions between subsystems and components within one system. These standards include possible mechanisms to store data securely and protect data as it's exchanged between subsystems and components. They specify the need for architecture and operations necessary to identify additional standards required to support multi-vendor systems and applications.

3.2 Formats for Data Interchange

Standards of data interchange formats are mode specific and specify meaning, representation and content of formats. These standards are used for the interchange of data between multiple systems. They identify specific formats for transfer and notation that separate transfer syntax from content definition, providing independent platforms. There are modality specific formats for biometric systems such as Finger Minutiae Format for Interchange, Face Recognition Format for Data Interchange, Iris Interchange Format and Finger Image Based Interchange Format.

3.3 Standards of Application Profiles

The standards of application profiles are developed to enable interoperability of information within system applications. These are specifications for at least one base standard relative to other standardized profiles. When necessary, they identify options, parameters, chosen classes and conforming subsets of base standards and other relevant profiles. These special applications of these standards are Verification of Transportation Workers and Border Management, as well as Point of Sale applications.

3.4 Standards for Testing and Reporting

Standards of performance testing and reporting define testing methodology and the requirements of reporting test results. These standards specify methods of calculations, definitions of metrics used, testing protocols and scenario testing protocols.

IV. THE DEVELOPERS OF BIOMETRIC STANDARDS

The Biometric standards are developed by the government agencies and the standard development organizations (SDOs). The committees and institutes who are involved in development of biometric standards are;

- i. National Institute of Standards and Technology (NIST)
- ii. International Committee for Information Technology Standards (INCITS) M1
- iii. Joint Technical Committee 1 (JTC1)/Subcommittee 37(SC37)
- iv. Organization for the Advancement of Structured Information Standards (OASIS)

The brief description of standard development organizations is included in table-1. This table contains the names of SDO, The basic information about SDO, The tasks these SDOs deal with, and the deliverables of each SDO. All this information is extracted from [5][6].

Table 1. A brief description of each SDO .

SDO	About SDO	Deals with	Standard Task Group/ Works groups/Deliverables
NIST	Secretary of Commerce Under information of technical management reform act.	Standards and Guidelines for Federal Computer Systems. Specially handles security and interoperability.	<i>DELIVERABLES</i> NISTIR- NIST interagency report- NISTIR6529-A (CBEFF) NIST special Publications SP-500-245 ANSI/NIST-ITL 1-2000 data formats for interchange of fingerprint, facial & scar mark and tattoo information.
INCITS M1 (Membership Open to any Organization) It also serves as	Established in November 2001 by ANSI	Information and communication Technology (ITC), Store, Process, Transfer,	<i>STANDARD TASK GROUPS</i> M1.2 (Biometric technical Interfaces) M1.3 (Biometric Data Interchange Formats)

technical Advisory Group(TAG) for ISO/IEC (JTC1)		Display and manage information	M1.4 (Biometric profiles) M1.5 Biometric performance Testing and Reporting) M1.6 (Societal Aspects of Biometric implementation)
JTC1/ SC 37 (Has twenty-one participating countries, six observer countries, and eleven liaison organizations)	Established in June 2002, By by JTC1	The international standardization projects for generic biometric technologies to support data interchange, interoperability, and testing	<i>WORKING GROUPS</i> WG1-Harmonized Biometric Vocabulary WG2- Biometric Technical Interfaces(standards for BioAPI and CBEFF) WG3- Biometric Data Interchange Formats WG4- Biometric Functional Architecture and related profiles WG5-Biometric Testing and Reporting WG6-Cross jurisdictional and societal aspects
OASIS (Not-for-profit, international consortium)	Founded in 1993 (has more than 5000 participants, from 600 organizations and 100 countries)	Web services standards for security, e-business and standardization efforts in the public sector and for application specific market.	OASIS XCBF- Standard way to describe identity OASIS TC defined a XML encodings based on ASN.1 schema of ANSI X.84:2003 They confirm XER-XML encoding rules. SAML- Security Assertion Markup Language developed by OASIS-SSTC

V. FACE BIOMETRIC STANDARDS

Face recognition is the automated process for recognizing individual by using facial characteristics. It is considered as the most challenging biometric method because of the perceptually identical structure of human faces, the possible environmental changes and variance in image capture conditions. ISO standards 19794-5 is the international standard for face recognition [7], it describes interchange formats for several types of biometric data. ISO/IEC 19794-5 defines specifically a standard scheme for codifying data describing human faces within a CBEFF-compliant data structure for use in facial recognition systems; it implements interoperability among vendors [8]. This international standard is intended to provide a face image format for face recognition applications requiring exchange of face image data. It has significant impact on both government and civilian biometric implementation such as E-passport, personal identity documents and access control systems.

The four main requirements in ISO 19794-5 are:

Photographic Specifications.

This standard specifies the photographic conditions such as lighting conditions, positioning of cameras as well as focus set by cameras.

Scene Constraint Specifications.

This standard specifies the scene constraints such as pose and expressions.

Digital Specifications

It also specifies the digital image attributes so as to have standard image resolution and image size. It helps to verify some identification marks and will also work on the devices with limited memory space.

Image Data Format Specifications.

The face image data format should follow specific format provided by the ISO standard to specify all details.

These four requirements help in improving face recognition accuracy.

VI. THE FACE IMAGE TYPES AND RECORD FORMAT

The ISO/IEC 19794-5 standards provide four face image types these types are Basic, Frontal, Full Frontal and Token Frontal. The Basic image is the fundamental face image type which specifies a record format including header and image data. There are no mandatory scenic, photographic and digital requirements for this image type. The Frontal is a kind of basic face image type that adheres to additional requirements appropriate for frontal face recognition. The frontal images are also used for human examination. The fundamental requirements are the limitation on head rotation angle, shadows over the face, subject lighting, scene lighting. Another type is Full Frontal type. In this type frontal images are specified with sufficient resolution for human examination as well as reliable automated face recognition. These types of images also include the full head as well as the neck and shoulders. These images are used to store the facial information permanently. Token Frontal is the fourth face image type which specifies frontal images with a specific geometric size. The main specification is regarding eye positioning based on the width and height of the image. The most important use of this type of image is to minimise the storage requirements on automated face recognition.

Another important specification is regarding image record format. The detail implementation of this format is as shown in the Figure 3 which is extracted from ISO/IEC 19794-5: Biometric Data Interface Format - Face Image Data [5]. This organization of the record format includes fixed-length (14 byte) *Facial Record Header* containing information about the overall record, including the number of facial images represented and the overall record length in bytes and a *Facial Record Data block* for each facial image

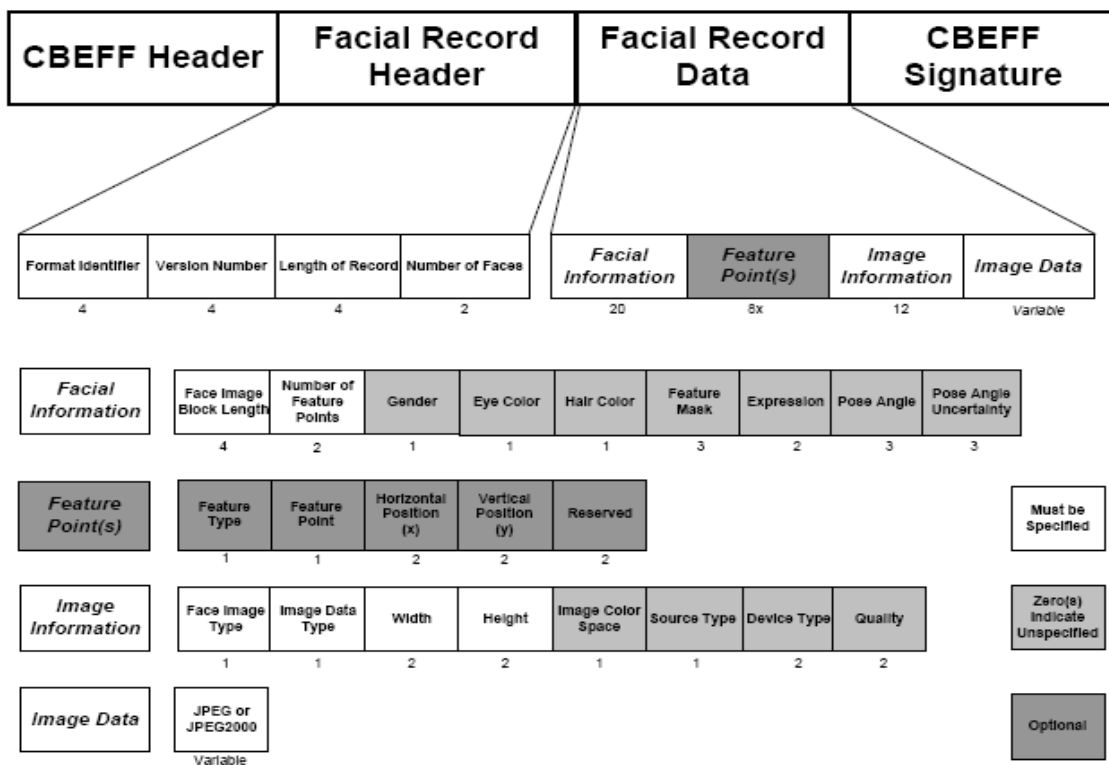


Figure 3: Face Image Record Format source: ISO/IEC 19794-5

This data consists of fixed length (20 byte) *Facial Information block* describing discernable features of the subject such as gender, multiple (including none) fixed length (8 byte) feature points block describing feature points in a facial image, A fixed length (12 byte) *Image Information block* describing digital features of the image such as face image type and dimensions such as width and height.

The standard provides the normative requirements for all the four types of face images; basic, frontal, full frontal and token frontal. These image-type specific requirements includes expected specification

of facial image like its encoding format, degree of rotation, camera position, lighting condition, resolution of the image. The face image provider and vendors should ensure that the normative requirements for the chosen image type are met [9]. While face image captured is normally handled by third party, it is logical for the biometric system integrator or biometric solution provider to implement automated conformance testing of ISO/IEC 19794-5 to ensure that input face images are compliant with the respective requirements of face image types.

By establishing standard formats for facial images the following objectives can be achieved[11]:

- . Allow interoperability among facial recognition vendors
- . Minimize the amount of data necessary to store face information with applications that have limited storage
- . Ensure that enrolled images will meet a quality standard needed for face recognition
- . Improve system throughput by saving the intermediate data instead of the raw data

VII. THE END-USER ADOPTION OF THE BIOMETRIC STANDARDS.

Standards are useful only if they are adopted. There is generally a lag time between the availability of standards and the availability of compliant products [10]. Further, many times vendors delay implementing the standards until they see customer demand for compliance. Some of the examples of end-user adoption of standards are listed below.

Table 2 . Examples of end-user adoptions

End user adoptions	Application
E-Passports	The International Civil Aviation Organization (ICAO) of the UN sets the requirements for machine readable travel documents (MRTDs), including e-passports and visas. ICAO has required that the biometrics stored within the e-passport conform to the requirements of the SC37 biometric data interchange format for face, fingerprint, and iris data.
Seafarer Identification.	The International Labour Organization (ILO) of the UN has a program for issuing a common identification credential for seafarers. This program has required that the fingerprint minutiae templates stored on the seafarer ID card conform to ISO/IEC 19794-2.
US Department of Homeland Security	DHS has required the use of INCITS biometric standards in several of its large biometric projects to include: <ul style="list-style-type: none"> • US Visitor and Immigration Status Indicator Technology (US VISIT) border management program • Transportation Worker Identification Credential (TWIC) • TSA Registered Traveler program
US Department of Defense	A number of INCITS standards have been adopted within the DoD Joint Technical Architecture and the Defense Information Standards Registry.
US Federal Employee Personal Identity Verification.	To comply with Homeland Security Presidential Directive (HSPD) 12, NIST developed technical specifications for the associated biometric-based credentialing system. Included in these specifications are requirements for compliance with the INCITS biometric data format specifications for finger images, minutiae templates, and facial images. Product availability is in progress, particularly since most of the standards are so recent, but a good example is the availability of BioAPI compliant products. BioAPI was released in 2001 and became an official ANSI standard in 2002. At this point, approximately 40 products have been announced

VIII. CONCLUSION

Biometric recognition can be described as automated methods to accurately recognize individuals based on distinguishing physiological and/or behavioural traits. The biometric traits plays vital role in making the biometric systems robust and performance efficient it is an increasingly critical component in the protection of information, infrastructure and personal identity, the continued development of comprehensive biometric standards is essential to ensure reliability, security, interoperability, usability and scalability. The International biometric standards are being widely used in both government and civilian applications. Especially the face recognition standards help in improving performance accuracy for identity and passport verification. Besides making the biometric system interoperable it also help in reducing the computational cost because of the less requirement of image pre-processing and image registration task. The Image database containing the images in standard format improves efficiency, robustness and interoperability of the biometric system.

REFERENCES

- [1] Biometrics.gov: Biometrics history (2006).
<http://www.biometrics.gov/Documents/BioHistory.pdf>
- [2] Biometrics and its application, BioEnable technology private Ltd.
www.bioenabletech.com
- [3] Emilio Mordini and Sonia Massari (2009) , Body, Biometrics and Identity, *Bioethics* ISSN 0269-9702 (print); 1467-8519 (online) doi:10.1111/j.1467-8519.2008.00700.x Volume 22 Number 9
- [4] Mr. Enji Hutchinson (2008), Biometric Standards for DoD Operational Requirements, *Biometrics task force*
- [5] NISTC-Biometric standards.
- [6] Fernando Podio, Overview of National and International Biometric Standards Activities *NIST Biometric Standards Program Computer Security Division, NIST/ITL*
- [7] ISO/IEC 19794-5: Biometric Data Interface Format - Face Image Data.
- [8] Lim Eyung, Lum Jia Jun Brandon, Dai Zhong Min, Face Biometric Standards and Conformance, Temasek Polytechnic/Cyber and Digital Security
- [9] NISTC Policy for Enabling the Development, Adoption and Use of Biometric Standards (2007), *NISTC Subcommittee on Biometrics and Identity Management*.
- [10] Creed Jones, et al , Description of Biometric Data Interchange Format Standards, INCITS M1 Technical Editors
- [11] Cathy Tilton, Biometric Standards – An Overview, January 2006

Authors

N. M Thakare is an Assistant Professor, working in the department of computer Science and Engg at Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Amravati, Maharashtra (India). She has completed ME in Computer Science and Engineering. Currently she is pursuing Ph.D. under the supervision of Dr. V. M. Thakare. Her Area of research is Computer Vision. She has published and presented 17 papers at National and International level. Currently she is working on pattern recognition algorithms and the face recognition methodologies



V M Thakare is Professor and Head of PG department of computer Science and Engg in SGB Amravati University Amravati, Maharashtra (India) and has completed ME in Advance Electronics and Ph.D. in computer Science/Engg. His Area of Ph D is Robotics and Artificial Intelligence. Currently he is working in area of wireless computing, mobile computing, Information Technology. He is Recognized supervisor for computer science and computer engineering in this University and also in other universities. He has also received national level Award for excellent paper. More than 10 candidates are working for Ph D Under his supervision. He has published and presented more than 115 papers at National and international level.



COMPARATIVE ANALYSIS OF LOW-LATENCY ON DIFFERENT BANDWIDTH AND GEOGRAPHICAL LOCATIONS WHILE USING CLOUD BASED APPLICATIONS

N. Ajith Singh¹ and M. Hemalatha²

¹Ph.D Research Scholar, Department of Computer Science, Karpagam University

²Head, Dept of Software Systems, Karpagam University, Coimbatore – 21

ABSTRACT

Cloud computing, a approach of computing where scalable and elastic IT-related capabilities are provided as shared assorted services (IaaS, PaaS,SaaS, DaaS), metered by use, to customers using internet technologies built on top of diverse technologies like virtualisation, distributed computing, utility computing, and more recently networking, web infrastructure and software services. It represents a paradigm shift in how we think about our data, the role of our computing devices and on managing computing resources. Cloud computing comes with Zero latency so the application hosted on cloud computing will act just like desktop application. The speed of internet is main factor to avoid latency but internet speed is not same in all the geographical location. From the survey it's found out that some latency occurs in different geographical location. Even different browser gives different latency. The problem of latency or low bandwidth may be automatically solve in coming years as the speed of internet is changing day to day. If users are going to adopt cloud computing then user must think of internet speed which is available in mean time. The internet connection should be good enough to support cloud computing minimum requirement would be 1MPBS.

KEYWORDS: Cloud computing, bandwidth, latency, web applications

I. INTRODUCTION

Outsourcing of data center functionality and availability of desktop application online in web mode via a network connection is what we tem cloud computing [50][52][51]. Companies and SME are all under the cloud computing[49][48] many businesses has been move to cloud, moving of cloud cut down the IT cost having the security with less IT staff, having the power of scalability and elasticity of cloud computing has changes who we use the computer. The expenses have been cut down and traffic of network has increase to double times [47]. Having a big question in mind the future of cloud service will replace desktop application and we all will depend on internet [15][19]. Will the future of internet support us? A concern for many organizations is that cloud computing relies on the availability, quality and performance of their internet connection. Poor application performance causes companies to lose customers, reduce employee productivity, and reduce bottom line revenue [40][42]. Because application performance can vary significantly based on delivery environment, businesses must make certain that application performance is optimized when written for deployment on the cloud or moved from a data center to a cloud computing infrastructure [41]. Dependence on an internet connection raises some key questions:

- What happen if we lose internet connection?
- What if the net connection which we use does not support SaaS application

Moving an existing in-house application or new application to the cloud will almost surely have some trade-offs in terms of performance [43][44][45]. Most existing enterprise applications won't have been designed with the cloud in mind [54][53]. Organizations considering moving to cloud computing will certainly have to think in costs for improving the network infrastructure required to run applications in the cloud [10][59][58][57][56]. On the other side, bandwidth continues to increase and approaches such as dynamic caching, compression, pre-fetching and other related web-acceleration technologies can effect in major performance improvements for end users, often exceeding 50% [60]. The application has to design or redesign in such a way that it will support cloud computing in order to achieve maximum performance [12][60]. Latency [55] (the time taken, or delay, for a packet of data to get from one designated point to another) will undoubtedly be an issue for certain applications. CRM or stock trading applications which require near-zero latency will probably be run in-house for many years to come [25]. Latency on the Internet is extremely variable and unpredictable there may a problem when putting these applications in the cloud [61][62][63]. There are many cloud computing commentators who claim that the cloud will never be able to support these types of applications. However, vendors such as Juniper and IBM are already demonstrating extremely low latency capabilities in the cloud so it's a case of watch this space [1][39].

Any desktop application which was being hosted in local server will be hosted in cloud services [26]. The power of desktop application may not be able to support so much power resource on network [27]. Cloud is providing preconfigured infrastructure at lower cost, which generally follows the Information Technology Infrastructure Library, can manage increased peak load capacity and moreover uses the latest technology, provide consistent performance that is monitored by the service provider [37]. Dynamic allocation of the resources as and when is needed. Cloud computing reduces capital expenditure and it offers high computing at lower cost [21][38]. Upgrading your hardware/software requirement also easy with the cloud, without disturbing the current work. Scalability and maintenance is easy in the case of cloud. Easily user can rent/lease the services offered by cloud computing vendors [2][13][14]. Latency may cause security problem while using cloud computing as the connection will be alive for long time [33]. SLA has to be taken care while adopting cloud computing by the user [34][35]. Security has to be taken in cloud storage also [36].

User will be charged as pay per usage like utility based services [31]. Overall Cloud is giving good performance at lower cost instead of making more capital investment [32]. The rest of the paper is organized as follows: Section 2 discusses about the Cloud providers. Section 3 Related work Section 4 Problem Definition and technical study, and Conclusion.

II. CLOUD PROVIDER

A. *Infrastructure as a Service (IaaS)* provisions hardware, software, and equipments to deliver software application environments with a resource usage-based pricing model. Infrastructure can scale up and down dynamically based on application resource needs [3].

B. *Platform as a Service (PaaS)* offers a high-level integrated environment to build, test, and deploy custom applications. Generally, developers will need to accept some restrictions on the type of software they can write in exchange for built-in application scalability. An example is Google's App Engine, which enables users to build Web applications on the same scalable systems that power Google applications, Web application frameworks, Python Django (Google App Engine), Ruby on Rails (Heroku), Web hosting (Mosso), Proprietary (Azure, Force.com) [3].

C. *Software as a Service (SaaS)* User buys a Subscription to some software product, but some or all of the data and codes resides remotely. Delivers special-purpose software that is remotely accessible by consumers through the Internet with a usage-based pricing model. In this model, applications could run entirely on the network, with the user interface living on a thin client. Salesforce is an industry leader in providing online CRM (Customer Relationship management) Services. Live Mesh from Microsoft allows files and folders to be shared and synchronized across multiple devices. Identity (OAuth, OpenID), Integration (Amazon Simple Queue Service), Mapping (Google Maps, Yahoo! Maps), Payments (Amazon Flexible Payments Service, Google Checkout, PayPal), Search (Alexa, Google Custom Search, Yahoo! BOSS), Others (Amazon Mechanical Turk) Other than the listed above companies, many companies started offering cloud computing services[3].

III. RELATED WORK

Cloud computing promise to deliver desktop like application in web which is support by any internet enabled device. Web applications are now hosted in elastic cloud environments where the unit of resource allocation is a virtual machine (VM) instance. A variety of techniques can reduce the latency of communication between VMs co-located on the same server in, say, a private cloud[17] on elastic cloud systems, there will be co-located VMs of the same web application on the same physical server that can take advantage of shared-memory IPC between the VMs. It is shown that Nahanni memcached, used with VDE networking, can improve the total read-related latency for a workload by up to 45% (i.e., read latest workload) compared to standard memcached, resulting from reductions in cache read latency of up to 86%. When combined with state-of-the-art paravirtualized network mechanisms, such as vhost, Nahanni memcached can still reduce the total read-related latency by up to 29%.[5][6][8][9][11]

IV. PROBLEM DEFINITION AND TECHNICAL STUDY

Latency seems to be major problem. Latency is also depends on geography and distance from server to host. Great latency happens when the end user and internal networks system is far from the cloud environment [20][21]. End user is platform dependent and all depend on net speed, if the end user is not having high speed connection then end user has to cost more money to get the cloud speed [5] [30]. A minimum requirement to use cloud is 1mpbs. All cloud is going to cost us in high bandwidth even if they provide us the all *.aas service [31][32]. We can see the difference of cloud when we access in cyber café or GPRS connection, when I check in cyber café and try to open Google docs it took 20 sec while when it is try to open at the campus of university which provide 5.4 mbps it open in 2 sec.

Below is a technical report of some cloud provider.

The below test is taken at www.cloudsleuth.net

City	Response Time ▲
Singapore	0.69 sec
Mumbai	4.04 sec
Beijing	4.96 sec
Tokyo - NTT	5.09 sec
Tokyo - KDDI	5.13 sec
Sydney	7.22 sec

Fig 1.Cloud Providers/ Windows Azure(Southeast Asia-Singapore) 5.4mbps

City	Response Time ▲
Singapore	1.13 sec
Mumbai	3.93 sec
Beijing	4.71 sec
Tokyo - NTT	5.06 sec
Tokyo - KDDI	5.21 sec
Sydney	7.15 sec

Fig 2. Cloud Providers/ Windows Azure(Southeast Asia-Singapore) From 56kbps

City	Response Time
Beijing	5.19 sec
Mumbai	12.82 sec
Sydney	13.34 sec
Tokyo - KDDI	13.45 sec
Tokyo - NTT	13.98 sec
Singapore	14.63 sec

Fig 3 Cloud Providers/ Google App Engine from 5.4 mbps

City	Response Time
Singapore	13.80 sec
Tokyo - KDDI	14.46 sec
Tokyo - NTT	14.96 sec
Sydney	15.89 sec
Mumbai	16.46 sec
Beijing	N/A

Fig 4 Cloud Providers/ Google App Engine from 56kbps

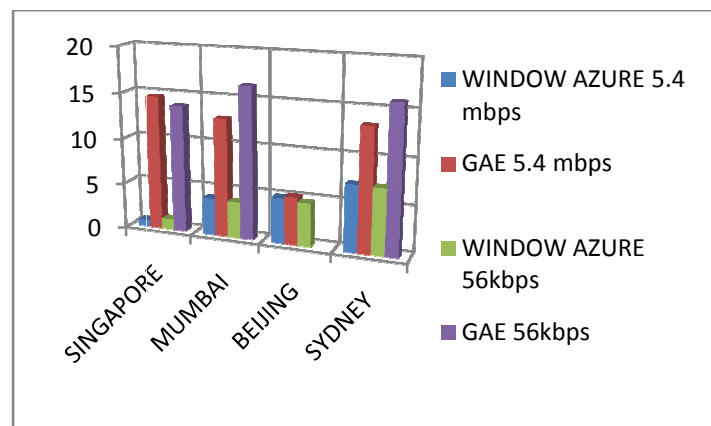


Fig 5 Graphical Represent of latency

Browser and Location Response Time Differences (seconds)*		
Location	Internet Explorer 7	Firefox 3.5
New York, NY	0.467	0.872
Seattle, WA	1.225	1.428
Los Angeles, CA	0.794	1.555
Chicago, IL	0.516	0.905

Fig 6 Location Response Time Differences

URL tested: <https://docs.google.com/?tab=mo&authuser=0&pli=1#home>

From the above test which was conducted on www.cloudsleuth.net in two major cloud application window azure and Google app engine, I have found out that in different city the response time is slight up and down I also observe that some site browser controls the latency of the web application. The test is taken with same net connection with 5.4 mbps. Browser latency was taken on google docs because we all use google docs. If the speed on high bandwidth 5.4mpbs some latency is there then I believe that there would be a big gap with different net connection. In India the net speed connection is not always same. Even BSNL who provides broad band connection to all nation of India don't provide same speed on different state. Speed of BSNL provided in northeast India and speed of BSNL provided at southern region is not same.

If all we depend on cloud then there would be huge gap among the user also. If the user want to utilize the cloud then he has spend more money on it. We have to take a best net connection to use cloud computing. But in recent years lots of changes has came from 2g to 3g and network service is much better than before but can cloud computing can adapt in India both rural and urban.

V. CONCLUSION

In the context of different geographic location and different web browser web application has to be design to avoid latency. Web application is design to give us like desktop application any delay in latency would be a great lose too many clients. Cloud computing expect to give us no latency but from the survey it is found that latency occurs on different geographical location. Above diagram and chart shows that difference in response time in different location with different web browser. A better algorithm is needed to prevent the latency while using the cloud computing. Cloud computing represents a powerful and proven solution for IT departments to increase elasticity and cut costs typically associated with the deployment of new platforms. However, the consideration of cloud computing should not be limited to this view and should take into consideration the entire system solution – from Cloud to end-user. Use of redundant cloud architecture, redundant data center infrastructure and an innovative route optimization technology can provide the required performance and availability of applications. Additionally, use of technologies beyond the cloud edge such as TCP acceleration and content delivery networks can further improve content delivery to end-users.

ACKNOWLEDGEMENT

We thank Karpagam University for motivating and encouraging doing our Research work in a Successful

REFERENCES

- [1] <http://www.docstoc.com/docs/40593418/Cloud-Computing---A-Management-Guide>
- [2] Enhancing Cloud Security through Policy Monitoring Techniques, V.V. Das and N. Thankachan (Eds.): CIIT 2011, CCIS 250 pp. 172–177, 2011. © Springer-Verlag Berlin Heidelberg 2011
- [3] Cloud Computing Resource Management for Indian E-governance, V.V. Das and N. Thankachan (Eds.): CIIT 2011, CCIS 250 pp. 278–281, 2011.
- [4] A COMPREHENSIVE SOLUTION TO CLOUD TRAFFIC TRIBULATIONS, International Journal on Web Service Computing (IJWSC), Vol.1, No.2, December 2010
- [5] N. Chohan, C. Bunch, and S. Pang. AppScale: Scalable and Open AppEngine Application Development and Deployment. In First International Conference on Cloud Computing, Munich, 2009.
- [6] Low-Latency Caching for Cloud-Based Web Applications, NetDB'11, June 12, 2011, Athens, Greece. Copyright 2011 ACM 978-1-4503-0652-2/11/06
- [7] Google App Engine. <http://code.google.com/appengine/>.
- [8] Memcached - a distributed memory object caching system. <http://memcached.org/>.
- [9] Nahanni. <http://gitorious.org/nahanni>.
- [10] Openstack open source cloud computing software. <http://openstack.org/>.

- [11] Twitter engineering: Memcached SPOF mystery. <http://engineering.twitter.com/2010/04/memcached-spoof-mystery.html>.
- [12] Windows azure. <http://www.microsoft.com/windowsazure/>.
- [13] J. Appavoo, A. Waterland, D. Da Silva, V. Uhlig, B. Rosenberg, E. Van Hensbergen, J. Stoess, R. Wisniewski, and U. Steinberg. Providing a cloud network infrastructure on a supercomputer. In Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, pages 385–394, Chicago, Ill, 2010. ACM Press.
- [14] E. Bugnion, S. Devine, and M. Rosenblum. Disco: running commodity operating systems on scalable multiprocessors. ACM SIGOPS Operating Systems Review, 31(5):143–156, Dec. 1997.
- [15] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber. Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems (TOCS), 26(2), 2008.
- [16] R. Grossman. The Case for Cloud Computing. IT Professional, 11(2):23–27, Mar. 2009.
- [17] F. J. Krauthem. Private virtual infrastructure for cloud computing. In Hot Topics in Cloud Computing (HotCloud 2009). USENIX Association, June 2009.
- [18] A. Lakshman and P. Malik. Cassandra: a decentralized structured storage system. ACM SIGOPS Operating Systems Review, 44(2):35, Apr. 2010.
- [19] G. Lee, N. Tolia, P. Ranganathan, and R. H. Katz. Topology-aware resource allocation for data-intensive workloads. In Proceedings of the first ACM Asia-Pacific workshop on systems - APSys '10, New Delhi, India, Aug. 2010. ACM Press.
- [20] C. Macdonell. Fast Shared Memory-based Inter-Virtual Machine Communications. PhD thesis, University of Alberta, 2011.
- [21] X. Meng, V. Pappas, and L. Zhang. Improving the Scalability of Data Center Networks with Traffic-aware Virtual Machine Placement. In 2010 Proceedings IEEE INFOCOM, pages 1–9. IEEE, Mar. 2010.
- [22] L. Nealan. Caching & performance: Lessons from Facebook. OSCon 2008, <http://www.scribd.com/doc/4069180/Caching-Performance-Lessons-from-Facebook>, July 2008.
- [23] D. Nurmi, R. Wolski, C. Grzegorzczak, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov. The Eucalyptus Open-Source Cloud-Computing System. In 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, pages 124–131. IEEE, May 2009.
- [24] J. T. Piao and J. Yan. A Network-aware Virtual Machine Placement and Migration Approach in Cloud Computing. In 2010 Ninth International Conference on Grid and Cloud Computing, pages 87–92. IEEE, Nov. 2010.
- [25] Hyukho Kim, Hana Lee, Woongsup Kim, Yangwoo Kim, “A Trust Evaluation Model for QoS guarantee in Cloud Systems”, International Journal of Grid and Distributed Computing, Vol.3, No.1, March, 2010.
- [26] Ghalem Belalem, Samah Bouamama, Larbi Sekhri, “An Effective Economic Management of Resources in Cloud Computing”, JOURNAL OF COMPUTERS, VOL. 6, NO. 3, MARCH 2011, pg no: 404-411.
- [27] Aishwarya C.S. and Revathy.S, “Insight into Cloud Security issues”, UACEE International journal of Computer Science and its Applications, pg no: 30-33.
- [28] Amreen Khan and Kamal Kant Ahirwar, “MOBILE CLOUD COMPUTING AS A FUTURE OF MOBILE MULTIMEDIA DATABASE”, International Journal of Computer Science and Communication, Vol. 2, No. 1, January-June 2011, pp. 219-221.
- [29] Richard Chow, Philippe Golle, Markus Jakobsson, Ryusuke Masuoka, Jesus Molina Elaine Shi, Jessica Staddon, “Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control”, CCSW'09, November 13, 2009.
- [30] Muzafar Ahmad Bhat, Razeef Mohd Shah, Bashir Ahmad, “Cloud Computing: A solution to Geographical Information Systems (GIS)”, International Journal on Computer Science and Engineering (IJCE), Vol. 3 No. 2 Feb 2011, pg no: 594-600.
- [31] Toward publicly auditable secure cloud data storage services, Cong Wang Kui Ren Wenjing Lou Jin Li, Journal IEEE Network: The Magazine of Global Internetworking archive Volume 24 Issue 4, July-August 2010
- [32] BNA. Privacy & security law report, 8 PVLR 10, 03/09/2009. Copyright 2009 by The Bureau of National Affairs, Inc. (800-372-1033), 2009 <http://www.bna.com> [accessed on: 2 November 2009]
- [33] Kandukuri BR, Paturi VR, Rakshit A. Cloud security issues. In: IEEE international conference on services computing, 2009, p.517–20.
- [34] Softlayer. Service Level Agreement and Master Service Agreement, 2009 /<http://www.softlayer.com/sla.html> [accessed on: 11 December 2009]

- [35] Bowers KD, Juels A, Oprea A. HAIL: a high-availability and integrity layer for cloud storage, Cryptology ePrint Archive, Report 2008/489, 2008 /<http://eprint.iacr.org> [accessed on:18October2009]
- [36] Survey on Cloud Computing Security, Shilpashree Srinivasamurthy, David Q. Liu, 2nd IEEE International Conference on Cloud Computing, 2010
- [37] Security Guidance for Critical Areas of Focus in Cloud Computing, April 2009. DOI = <http://www.cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf>
- [38] Towards Analyzing Data Security Risks in Cloud Computing Environments Amit Sangroya, Saurabh Kumar, Jaideep Dhok, and Vasudeva Varma, Springer-Verlag Berlin Heidelberg 2010
- [39] Cloud Computing and Grid Computing 360-Degree Compared, Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu, Grid Computing Environments Workshop, 2008. GCE '08 2008
- [40] Management in an Application Data Cache. In Symposium on Operating Systems Design and Implementation (OSDI). USENIX Association, 2010.
- [41] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS '09), pages 199–212, Chicago, Ill, Nov. 2009. ACM Press
- [42] R. Russell. virtio: Towards a De-Facto Standard For Virtual I/O Devices. SIGOPS Oper. Syst. Rev., 42(5):95–103, 2008
- [43] P. Saab. Scaling memcached at Facebook. http://www.facebook.com/note.php?note_id=39391378919,2008
- [44] B. Sotomayor, R. S. Montero, I. M. Llorente, and I. Foster. Virtual Infrastructure Management in Private and Hybrid Clouds. IEEE Internet Computing, 13(5):14–22, Sept. 2009
- [45] A. Thusoo, Z. Shao, S. Anthony, D. Borthakur, N. Jain, J. Sen Sarma, R. Murthy, and H. Liu. Data warehousing and analytics infrastructure at Facebook. In Proceedings of the 2010 International Conference on Management of Data (SIGMOD '10), pages 1013–1020, Indianapolis, Indiana, June 2010. ACM Press
- [46] N. Tolia and M. Satyanarayanan. Consistency-preserving caching of dynamic database content. In International World Wide Web Conference, pages 311–320, Banff, Alberta, 2007
- [47] X. Zhang, S. McIntosh, P. Rohatgi, and J. L. Griffin. Xensocket: a high-throughput interdomain transport for virtual machines. In MIDDLEWARE2007: Proceedings of the 8th ACM/IFIP/USENIX international conference on Middleware, pages 184–203, Berlin, Heidelberg, 2007. Springer-Verlag
- [48] M. Armbrust, A. Fox, Griffith et al., “Above the clouds: A Berkeley view of cloud computing,” EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28, 2009
- [49] M. Turner, D. Budgen, and P. Brereton, “Turning software into a service,” Computer, pp. 38–44, 2003
- [50] M. Papazoglou, P. Traverso, S. Dustdar, and F. Leymann, “Service-oriented computing: State of the art and research challenges,” Computer, pp. 38–45, 2007
- [51] F. Chong and G. Carraro, “Architecture strategies for catching the long tail,” MSDN Library, Microsoft Corporation, 2006
- [52] F. Chong, G. Carraro, and R. Wolter, “Multi-Tenant Data Architecture,” MSDN Library, Microsoft Corporation, 2006
- [53] S. Aulbach, T. Grust, D. Jacobs, A. Kemper, and J. Rittinger, “Multi-tenant databases for software as a service: schemamapping techniques,” in Proceedings of the 2008 ACM SIGMOD international conference on Management of data, 2008, pp. 1195–1206
- [54] L. Tao, “Shifting paradigms with the application service provider model,” Computer, pp. 32–39, 2001
- [55] C. Guo, W. Sun, Y. Huang, Z. Wang, B. Gao, and B. IBM, “A framework for native multi-tenancy application development and management,” in International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007, pp. 551–558
- [56] “wso2.org/projects/carbon,” [Online; accessed 20-March-2010]
- [57] O. Alliance, “OSGi service platform, Core Specification release 4.1,” Draft, May, 2007
- [58] T. Andrews, Curbera et al., “Business process execution language for web services, version 1.1,” Standards proposal by BEA Systems, International Business Machines Corporation, and Microsoft Corporation, 2003
- [59] J. Kotamraju, “Java API for XML-Based Web Services (JAXWS 2.0),” 2006
- [60] S. Perera, C. Herath et al., “Axis2, middleware for next generation web services,” Proc. International Conference on Web Services, 2006, pp. 833–840, 2006
- [61] D. Jacobs and S. Aulbach, “Ruminations on multi-tenant databases,” BTW Proceedings, 2007
- [62] Dominik Birk, “Technical Challenges of Forensic Investigations in Cloud Computing Environments”, January 12, 2011

[63]Multi-Tenant SOA Middleware for Cloud Computing, Afkham Azeez, Srinath Perera, Dimuthu Gamage, Ruwan Linton, Prabath Siriwardana, Dimuthu Leelaratne, Sanjiva Weerawarana, Paul Fremantle WSO2 Inc.Mountain View, CA, USA

AUTHOR PROFILES

M. Hemalatha completed M.C.A M.Phil., Ph.D., in Computer Science and currently working as an Assistant Professor and Head, Department of software systems in Karpagam University. Ten years of Experience in teaching and published Seventy two papers in International Journals and also presented seventy papers in various National conferences and one international conference. Area of research is Data mining, Software Engineering, bioinformatics, Neural Network. Also reviewer in several National and International journals.



Ajith Singh. N is presently doing Ph.D., in Karpagam University, Coimbatore, Tamil Nadu, and India and has completed M.C.A degree in 2007 and B.Sc (Information Technology) in 2005. Major research area is Cloud Computing, Advance Networking. At Present, he is doing his research work under the guidance of Dr. M. Hemalatha HOD of Department of software systems in Karpagam University, Coimbatore.



EVALUATION OF TEXTURAL FEATURE EXTRACTION FROM GRLM FOR PROSTATE CANCER TRUS MEDICAL IMAGES

R.Manavalan¹ and K.Thangavel²

¹Department of Computer Science and Applications, KSR College of Arts and Science,
Tiruchengode, Namakkal, Tamilnadu, India

²Department of Computer Science, Periyar University, Salem 11

ABSTRACT

Ultrasound imaging is one of the promising techniques for early detection of prostate cancer. There are five steps involved in processing the ultrasound images such as pre-processing, segmentation, feature extraction, feature selection, and classification. In this paper, the Transrectal Ultrasound (TRUS) images are pre-processed with M3 filter. Then it is segmented by using DBSCAN clustering after applying morphological operators, in order to extort the prostate region. It is proposed to extract the features by using Gray Level Run Length Matrix (GLRLM) for different direction from the segmented region. To classify the images into benign or malignant, Support Vector Machine (SVM) is adapted to evaluate the performance of the proposed method through classification. Over 5500 digitized TRUS images of prostate are used for the experimental analysis. The results obtained using the classification showed that texture features based on GLRLM by using combined directions ($\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$) distinguish between malignant and benign on TRUS images, with highest accuracy 85% where as sensitivity and specificity of 82% and 100% respectively.

KEYWORDS: SVM, GLRLM, Prostate, DBSCAN, M3-Filter, Sensitivity, Specificity, Accuracy

I. INTRODUCTION

In Early detection of cancers such as prostate cancer, breast cancer, lungs cancer, cervical cancer and blood cancer, digital image plays a vital role. Prostate cancer is now most frequently diagnosed with one in every eleven men [19]. It is in the second position in cancer-related cause of death only for male [16]. To diagnosis and prognosis, Ultrasound imaging method is recommended in general. Because of the result of reflection, refraction and deflection of ultrasound waves from different types of tissues with different acoustic impedance, an accurate detection of region of interest in ultrasound image is crucial.

Normally, the contrast in ultrasound image is very low and boundary between region of interest and background are more uncertain [15]. Different types of noises affect the mages. The ultrasound image with speckle noise and weak edges make the identification of prostate region more difficult. Hence, the analysis of ultrasound image is more challenging one.

In General, there are two common use of ultrasound medical imaging: (i) To guide the oncologist in the biopsy procedure and (ii) To establish the volume of the prostate. This method is used in diagnosing for more than 50 years. To classify the proteomic into benign or malign, the statistical textural features are extracted from the segmented region. This paper emphasis the extraction of textural features using GRLM for various directions and Support Vector Machine (SVM) is adopted for classification. An overview of the proposed system is presented below.

II. OVERVIEW OF THE SYSTEM

There are five stages in the proposed system consists, such as acquisition of Transrectal Ultrasound (TRUS) image of prostate, pre-processing, segmentation, feature extraction and classification. For automatic detection of prostate tumour in TRUS image, Computer Aided Diagnosis (CAD) system is developed. It can provide the valuable viewpoint and accuracy of earlier prostate cancer detection. The overview of the proposed system is depicted in figure 1.

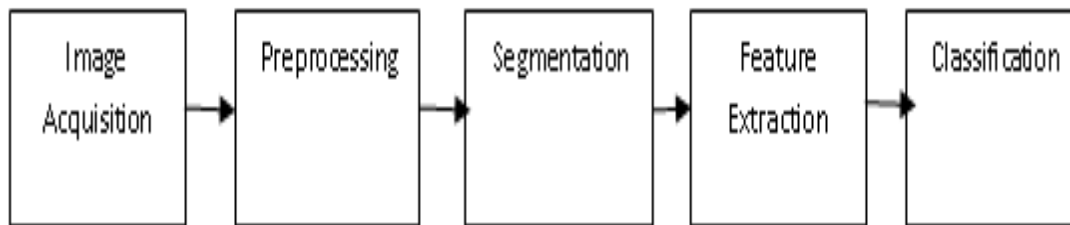


Figure 1: CAD System

The subsequent sections describe each stage in detail. The rest of the paper is organized as follows: section 2 deals about Image Acquisition and pre-processing of original TRUS medical image of prostate for noise removal and enhance the contrast of the image. The DBSACN clustering with morphological operators for locating the Region of Interest (ROI) from TRUS prostate medical image is described in section 3. The feature extraction through GLRLM for the classification of prostate cancer is illustrated in section 4. The SVM classifier is discussed in section 5. The experimental analysis and discussion are presented in the section 6. Section 7 concludes this work with further directions.

III. IMAGE ACQUISITION AND PREPROCESSING

Among the different medical imaging modalities, Ultrasound imaging is a widely used technology for prostate cancer diagnosis and prognosis [20]. Less than 30 minutes is required for each case of image acquisition processes. The TRUS images of the prostate gland are acquired while the TRUS probe is supported and manipulated by the TRUS Robot which uses joystick located next to the console. By rotating the probe about its axis with minimum displacement and deformation, the entire prostate is scanned. The surgeon can manually adjust the probe depth. Exact recording of images and corresponding TRUS frame coordinates were obtained. The gathered information is used in offline segmentation for the ultrasound image of the prostate gland. Via frame grabber, Image acquisition is done from ultrasound device. Using a series of TRUS images, image of the prostate can be generated [21]. Because of poor image contrast, speckle noise, and missing or diffuse boundaries in the TRUS images, the ultrasound images are very difficult to segment.

The speckle noise is commonly found in the ultrasound medical images and hence M3-Filter is used to generate a despeckled image [14]. Data dropout noise is common and this is generally referred to as speckle noise [18]. An error in data transmission causes this noise. The corrupted pixels are either set to the maximum value. The information of edges is needed for the proper segmentation. The imtophat filter is used to create required edges once the noise is removed to enhance the contrast of the image. The following section discusses the segmentation of this enhanced image using DBSCAN clustering.

IV. IMAGE SEGMENTATION

Enhanced images are changed into binary image where each pixel is restricted to a value of either 0 or 1 to simplify the image segmentation task. This process is accomplished by local adaptive thresholding method [22]. Morphological operators such as opening and closing are applied on these binary images [2, 9]. In TRUS medical images, isolating prostate from other white region in the threshold image is a difficult task. By using opening and closing operators, the thresholded image can be reconstructed with a large disk structuring element in order to isolate the object with high

likelihood to be part of the prostate [9]. Pixels of the thresholded images are grouped by using DBSCAN with the aim of separating background from TRUS image to target possible prostate [7, 17].

By clustering, it takes a binary (segmented) image, and delineates only significantly important regions. The outcome that expected is desired boundary of the TRUS prostate image. A sample image and its segmented image are provided in figure 2(a) and figure 2(b) respectively. The DBSCAN Algorithm for segmentation is detailed in figure 3.

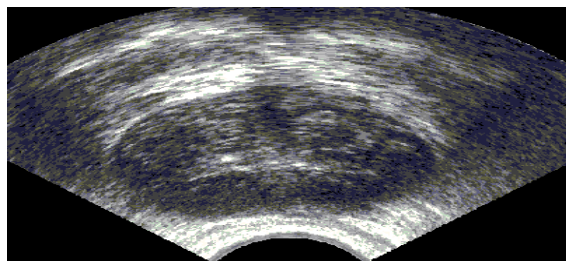


Figure 2(a) sample image



Figure 2(b) segmented image

DBSCAN Algorithm

INPUT: Enhanced TRUS prostate image

OUTPUT: Segmented image which contains only prostate

Step1: Set epsilon (eps) and minimum points (minPts).

Step2: Starts with an arbitrary starting point that has not been visited and then finds all the neighbor points within distance eps of the starting point.

Step3: If the number of neighbors is greater than or equal to minPts, a cluster is formed. The starting point and its neighbors are added to this cluster and the starting point is marked as visited. The algorithm then repeats the evaluation process for all the neighbors recursively.

Step4: If the number of neighbors is less than minPts, the point is marked as noise.

Figure 3: DBSCAN Algorithm

After fixing the two parameters eps and minPts, DBSCAN starts to cluster data points from an arbitrary point q [23]. It begins by finding the neighborhood of point q (all points that are directly density reachable from the point q). For an image, we start with left-top pixel (not necessarily a corner pixel, any arbitrary pixel can be chosen for first iteration) as our first point in the subimage. We look for first pixel satisfying the core pixel condition as a starting (seed) point. If the neighborhood is sparsely populated, i.e. it has fewer than MinPts points, and then the point q is labelled as noise [24]. Otherwise, a cluster is initiated and all points in neighborhood of the point q are marked by new cluster. Next the neighborhoods of the neighborhood q are examined iteratively to check if they can be added into the cluster [12]. If a cluster cannot be expanded further, DBSCAN chooses another arbitrary unlabeled point and repeats the process to form another cluster. This procedure is iterated till all pixels in the image have been identified as noise or with a cluster of prostate pixels.

V. FEATURE EXTRACTION

Texture analysis is normally grouped into four categories: model-based, statistical-based, structural-based, and transform-based methods. Model-based methods are based on the concept of predicting pixel values based on a mathematical model [14]. Statistical methods describe the image using numerical analysis of pixel intensity values. Transform methods generally perform some kind of

modification to the image, obtaining a new “response” image that is then analyzed as a representative proxy for the original images. Structural approaches seek to understand the hierarchical structure of the image. Texture is one of the important characteristics used in identifying an object in an image and to discriminate the images [1]. The texture coarseness or fineness of an image can be interpreted as the distribution of the elements in the matrix [10, 11]. In this paper, segmented image (ROI) is utilized to construct GLRLM using the various directions, and then seven features are extracted from each GLRLM for the classification of TRUS prostate image.

5.1. Gray Level Run Length Matrices (GLRLM)

The GLRLM is based on computing the number of gray-level runs of various lengths [3,4]. A gray-level run is a set of consecutive and collinear pixel points having the same gray level value. The length of the run is the number of pixel points in the run [3, 4]. The Gray Level Run Length matrix is constructed as follows:

$$R(\theta) = (g(i, j) | \theta), 0 \leq i \leq Ng, 0 \leq j \leq Rmax;$$

where Ng is the maximum gray level and Rmax is the maximum length. The following textural features are measured for the TRUS medical image of prostate. Let $p(i, j)$ be the number of times there is a run of length j having gray level i . There are five Run Length Matrix based features computed for 4 directions of run ($0^\circ, 45^\circ, 90^\circ, 135^\circ$, $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ run). $i=1, \dots, Ng$ and $j=1, \dots, Nr$.

Table 1: Features and their Formulae

S. No.	Features	Formulae
1	Short Run Emphasis (SRE)	$\frac{1}{n} \sum_{i,j} \frac{p(i, j)}{j^2}$
2	Long Run Emphasis (LRE)	$\frac{1}{n} \sum_{i,j} j^2 p(i, j)$
3	Grey Level Non-uniformity (GLN)	$\frac{1}{n} \sum_i \left(\sum_j p(i, j) \right)^2$
4	Run Length Non-uniformity	$\frac{1}{n} \sum_j \left(\sum_i p(i, j) \right)^2$
5	Run percentage (RP)	$\sum_{i,j} \frac{n}{p(i, j) j}$
6	Low Grey Level Run Emphasis (LGLRE)	$\frac{1}{n} \sum_{i,j} \frac{p(i, j)}{i^2}$
7	High Grey Level Run Emphasis (HGLRE)	$\frac{1}{n} \sum_{i,j} i^2 p(i, j)$

VI. SVM CLASSIFIER

Support vector machine is based on modern statistical learning theory. It is a useful method for classification. It is a leading method for solving non-linear problem. The SVMs were introduced by Vladimir Vapnik [25] and his colleagues. The earliest mention was in (Vapnik, 1979). Classification task usually involves training and testing data which consist of some data instances [5,6]. One target value and several attributes are contained into each instance of training set. Given the attributes of the testing set, the goal of SVM is to produce a model which predicts target value of data instances. Classification in SVM is an example of Supervised Learning. Whether the system is performing in the right way or not is indicated with the help of known labels. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A

step in SVM classification involves identification as which are intimately connected to the known classes [5].

SVM use the training data to create the optimal separating hyperplane between the classes. The optimal hyperplane maximizes the margin of the closest data points. A good separation is achieved by the hyperplane that has the largest distance to the nearest training features of any class (so-called functional margin). Maximum-margin hyperplane and margins for a SVM trained with samples from two classes. Samples on the margin are called the support vectors [8]. SVM divides the given data into decision surface. Decision surface divides the data into two classes like a hyper plane. Training points are the supporting vector which defines the hyper plane. The basic theme of SVM is to maximize the margins between two classes of the hyper plane (Steve, 1998).

Basically, Support Vector Machines are used to solve binary classification problems only. They are extended to handle multi-class problems also [6]. The idea is to decompose the problem into many binary-class problems and then combine them to obtain the prediction. To allow multi-class classification, SVM uses the one-against-one technique by fitting all binary sub classifiers and finding the correct class by a voting mechanism. The "one-against-one" approach (Knerr et al., 1990) is implemented in SVM for multiclass classification. If K is the number of classes, then $K(K - 1)/2$ binary classifiers are constructed and trained to separate each pair of classes against each other, and uses a majority voting scheme (max-win strategy) to determine the output prediction. For training data from the i th and the j th classes, we solve the following two-class classification problem [13].

Let us denote the function associated with the SVM model of $\{c_i, c_j\}$ as:

$$g(x)_{ij} = \text{sign}(f(x)_{ij})$$

An unseen example, x , is then classified as:

$$f(x) = \arg \max_i \sum_{i=1}^K \sum_{j=1 \wedge i \neq j}^K V_{ij}(x)$$

$$\text{where } V_{i,j}(x) = \begin{cases} 1 & \text{if } g_{ij}(x) = 1 \\ 0 & \text{if } g_{ij}(x) = -1 \end{cases}$$

Each feature set is examined using the Support Vector Machine classifier. In classification, we use a voting strategy: each binary classification is considered to be a voting where votes can be cast for all data points x in the end point is designated to be in a class with the maximum number of votes. In case that two classes have identical votes, though it may not be a good strategy, now we simply choose the class appearing first in the array of storing class names. The objective of any machine capable of learning is to achieve good generalization performance, given a finite amount of training data, by striking a balance between the goodness of fit attained on a given training dataset and the ability of the machine to achieve error-free recognition on other datasets. With this concept as the basis, support vector machines have proved to achieve good generalization performance with no prior knowledge of the data. The optimal separating hyperplane can be determined without any computations in the higher dimensional feature space by using kernel functions in the input space. Commonly used kernels include:

Linear Kernel: $K(x, y) = x \cdot y$

Radial Basis Function (Gaussian) Kernel: $K(x, y) = e^{(-\|x - y\|^2 / 2\sigma^2)}$

Polynomial Kernel: $K(x, y) = (x \cdot y + 1)^d$

VII. EXPERIMENTAL ANALYSIS AND DISCUSSION

The proposed method has been implemented using MATLAB. In order to evaluate this work, experiments conducted over 5500 image with normal and abnormal cases. The total of 5500 TRUS medical images of prostate measuring 250x250 pixels is transformed into a Grey Level Run Length matrix (GLRLM) along the directions of $(0^\circ, 45^\circ, 90^\circ, 135^\circ, \{0^\circ, 45^\circ, 90^\circ, 135^\circ\})$ producing 5 run length matrices. From each GLRLM, 7 texture features are extracted. Extracted features are fed into SVM to construct the prediction model. The performance of the prediction was evaluated in terms of sensitivity, specificity, accuracy, the respective formule are given in Table 2.

Table 2.Formula for Measures

Measures	Formula
Sensitivity	$TP/(TP+FN)$
Specificity	$TN/(TN+FP)$
Accuracy	$(TP+TN)/(TP+FP+TN+FN)$

Accuracy measures the quality of the classification. It takes into account true and false positives and negatives. Accuracy is generally regarded with balanced measure whereas sensitivity deals with only positive cases and specificity deals with only negative cases. TP is number of true positives, FP is number of false positives, TN is number of true negatives and FN is number of false negatives. A confusion matrix provides information about actual and predicted cases produced by classification system. The performance of the system is examined by demonstrating correct and incorrect patterns. They are defined as confusion matrix in Table 3.

Table 3.Confusion Matrix

Actual	Predicted	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

TP - predicts cancer as cancer.

FP - predicts cancer as normal.

TN - predicts normal as normal.

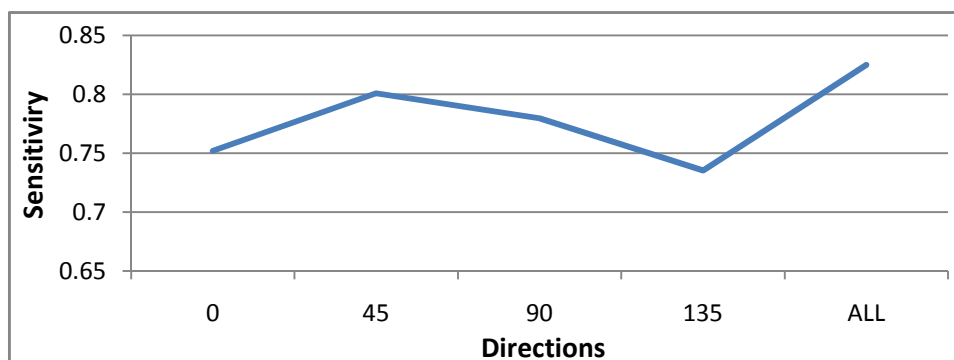
FN - predicts normal as cancer.

The higher value of both sensitivity and specificity shows better performance of the system. In order to quantify the influence of the direction, the GLRLM is constructed with direction choices for feature extraction. The results of performance measures for each feature set have been tabulated in table 4.

Table 4: results of statistical parameters for various directions

Direction	sensitivity	Specificity	Accuracy
GLRLM_000	0.75188	1	0.82945
GLRLM_045	0.800901	1	0.816667
GLRLM_090	0.779528	0.997886	0.821667
GLRLM_135	0.735294	1	0.814
GLRLM_ALL	0.824956	1	0.85

From the computational results we can notice that the features extracted from among GLRLM, the combined direction $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ give better results using three quantitative performance measures with higher sensitivity and specificity rate. The results of individual performance measures for various directions are depicted in the figure 5(a) – 5(c).

**Figure 5(a):** performance of sensitivity for various directions

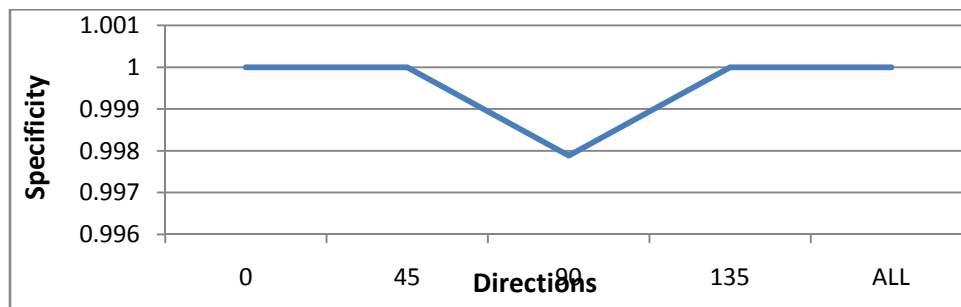


Figure 5(b): specificity for various directions

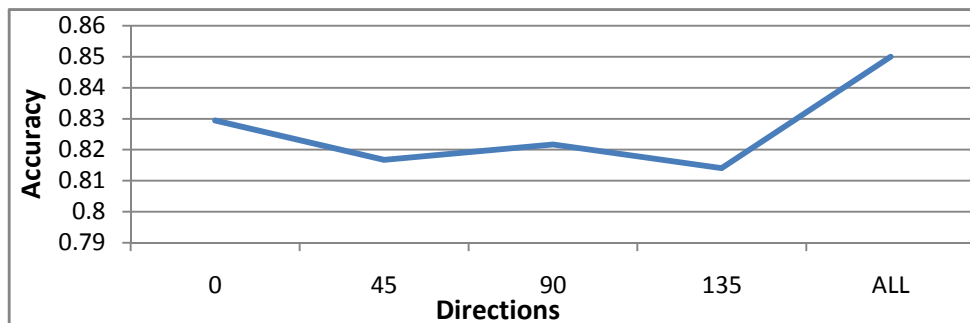


Figure 5(c): performance of accuracy for various directions

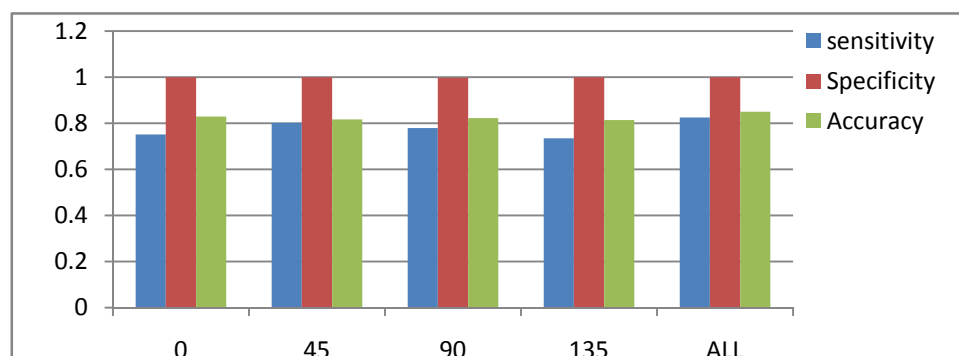


Figure 6: Relative performance measures for various directions

The performance of relative performance measures is depicted in figure 6 and observed that the features extracted from the combination of all directions ($\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$) producing the best one from the result of classification. It achieves 85% of the accuracy, 82% of the sensitivity and 100% of the specificity.

VIII. CONCLUSION

TRUS imaging is one of the best methods for detecting prostate cancer. In some cases, radiologists face difficulties in directing the tumours. To assist the oncologist and improve the accuracy of detection, the method presented in this paper. The proposed approach effectively addresses the feature extraction problem TRUS image analysis of prostate and can also be applied to other image analysis applications. The proposed method of feature extraction from GLRLM with various directions are compared and analyzed. The performance of the texture is measured using quantitative performance measures. The performance of proposed method is tested with ultrasound image regard to prostate. The computational results showed that texture features based on GLRLM by using combined directions ($\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$) distinguish between malignant and benign on TRUS images, with accuracy levels higher than texture features based on others. The maximum accuracy rate for normal and cancer classification is achieved 85%. The GLRLM features combined with other statistical

features to improve the results of classification of TRUS prostate images may be considered for future work.

REFERENCES

- [1] Amadasun, M. and King, R., (1989) 'Textural features corresponding to textural properties', IEEE Transactions on Systems, Man, and Cybernetics, vol. 19, no. 5, pp. 1264 - 1274.
- [2] Anil K. Jain. (1989) Fundamentals of Digital Image Processing: Prentice Hall.
- [3] K.Thangavel, M.Karnan, R.Sivakumar and A. Kaja Mohideen " Ant Colony System for Segmentation and Classification of Microcalcification in Mammograms" AIML Journal, Volume (5), Issue (3), pp.29-40., September, 2005.
- [4] Jong Kook Kim, Jeong Mi Park, Koun Sik Song and Hyun Wook Park. "Texture Analysis and Artificial Neural Network for Detection of Clustered Microcalcifications on Mammograms" IEEE, pp.199 – 206, 1997.
- [5] Devendran V et. al.,(2008) 'Texture based Scene Categorization using Artificial Neural Networks and Support Vector Machines: A Comparative Study,' ICGST-GVIP, Vol. 8, Issue IV, pp. 45-52.
- [6] Duda, R., Hart, P. And Stork, D. (2001), 'Pattern Classification': John Wiley and Sons.
- [7] Ester, M., Kriegel, H.P., Sander, J., and Xu, X.(1996) 'A density-based algorithm for discovering clusters in large spatial databases with noise'. Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, Portland: AAAI Press. pp. 226-231.
- [8] Fatima Eddaoudi, and Fakhita Regragui.:(2011) 'Masses Detection Using SVM Classifier Based on Textures Analysis' Applied Mathematical Sciences, Vol. 5, No. 8, pp. 367 – 379.
- [9] Gonzalez, R. and Woods, R. (2002). Digital Image Processing. 3rd Edn., Prentice Hall Publications, pp. 50-51.
- [10] Haralick, R.M. , Shanmugan, K.. and Dinstein, I.(1973) 'Textural Features for Image Classification', IEEE Tr. on Systems, Man, and Cybernetics, Vol SMC-3, No. 6, pp. 610-621.
- [11] Haralick, R.M. (1979) 'Statistical and Structural Approaches to Texture', Proceedings of the IEEE, Vol. 67, No. 5, pp. 786-804.
- [12] Massimiliano Pavan and Marcello Pelillo: (2007) 'Dominant Sets and Pairwise Clustering', IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 29, No. 1, pp. 167 – 172.
- [13] Metehan Makinaci:(2005) ' Support Vector Machine Approach for Classification of Cancerous Prostate Regions', World Academy of Science, Engineering and Technology. pp. 166 – 169.
- [14] Mohamed, S.S. and Salama M.M. (2005) 'Computer Aided diagnosis for Prostate cancer using Support Vector Machine' Publication: Proc., medical imaging conference, California, SPIE Vol. 5744, pp. 899-907.
- [15] Thangavel, K., Manavalan, R. and Laurence Aroquiaraj . I. (2009), Removal of Speckle Noise from Ultrasound Medical Image based on Special Filters: Comparative Study, ICGST-GVIP Journal, ISSN 1687-398X, Volume (9), Issue (III), pp. 25-32.
- [16] Scardino, PT.,(1989) 'Early detection of prostate cancer'. ; Urol Clin North Am., Vol. No 16(4) , pp: 635-55.
- [17] Sinan Kockara, Mutlu Mete, Bernard Chen, and Kemal Aydin:(2010). 'Analysis of density based and fuzzy c-means clustering methods on lesion border extraction in dermoscopy images', From Seventh Annual MCBIOS Conference: Bioinformatics Systems, Biology, Informatics and Computation Jonesboro, AR, USA. pp. 19 - 20.
- [18] Sinha, G.R., Kavita Thakur and Kowar, M.K. (2008). "Speckle reduction in Ultrasound Image processing", Journal of Acoustical. Society of India, Vol. 35, No. 1, pp. 36-39.
- [19] Yanong Zhu, Stuart Williams, and Reyer Zwiggelaar (2006), 'Computer Technology in Detection and Staging of Prostate Carcinoma: A review, Medical image Analysis' Vol. No.10, pp 178-199
- [20] Cancer Facts and Figures. American Cancer Society. [Online] <http://www.cancer.org>, 2002.

- [21] Huber, J.S. Qiyu Peng Moses, W.W. Reutter, B.W. Pouliot, J. Hsu, I.C (2011). “Development of a PET-Transrectal Ultrasound Prostate Imaging System” Nuclear Science, IEEE Transactions pp. 674 – 681.
- [22] Xiao-Ping Zhang, and Mita D. Desai.(2001) “Segmentation of Bright Targets Using Wavelets and Adaptive Thresholding” , IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 10, NO. 7, pp 1020 – 1030
- [23] Sait Suer, Sinan Kockara and Mutlu Mete(2011). “An improved border detection in dermoscopy images for density based clustering” biomedcentral, (Suppl 10):S12 pp. 1-10.
- [24] Xin Wang and Howard J. Hamilton DBRS (2003) “ A Density-Based Spatial Clustering Method with Random Sampling” Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science, Vol 2637/2003, 569.
- [25] Muntean, M. Vălean, H. Ileană, I. and Rotar, C.(2010). “Improving classification with Support Vector Machine for unbalanced data” Automation Quality and Testing Robotics (AQTR), 2010 IEEE International Conference ,Vol.3 pp. 1 – 6.

Authors

R. MANAVALAN Obtained M.Sc.,Computer Science from St.Joseph’s College of Bharathidasan University,Trichy, Tamilnadu, India, in the year 1999, and M.Phil., in Computer Science from Manonmaniam Sundaranar University, Thirunelveli, Tamilnadu, India in the year 2002. He works as Asst.Prof & Head, Department of Computer Science and Applications, KSR College of Arts and Science, Thiruchengode, Nammakal, Tamilnadu, India. He pursues Ph.D in Medical Image Processing. His areas of interest are Medical image processing and analysis, soft computing, pattern recognition and Theory of Computation.



THANGAVEL KUTTIANNAN received the Master of Science from Department of Mathematics, Bharathidasan University in 1986, and Master of Computer Applications Degree from Madurai Kamaraj University, India in 2001. He obtained his Ph. D. Degree from the Department of Mathematics, Gandhigram Rural University in 1999. Currently he is working as Professor and Head, Department of Computer Science, Periyar University, Salem. His areas of interests includes medical image processing, artificial intelligence, neural network, fuzzy logic, data mining, pattern recognition and mobile computing



ANALYSIS AND MULTINOMIAL LOGISTIC REGRESSION MODELLING OF WORK STRESS IN MANUFACTURING INDUSTRIES IN KERALA, INDIA

K. Satheesh Kumar¹ and G. Madhu²

¹Assistant Professor (Sr. Grade), Department of Mechanical Engineering, Federal Institute of Science and Technology, Ernakulam, Kerala, India.

²Professor and Head, Safety and Fire Engineering Division, Cochin University of Science and Technology, Ernakulam, Kerala, India.

ABSTRACT

This study examines the influence of factors responsible for work stress among the employees in the manufacturing industries in Kerala, India. The sample size of the subjects selected for the study consists of 75 Engineers, 110 Supervisors and 675 Workers in the selected manufacturing industries in Kerala, India. Seven factors were identified with the existing literatures, and in consultation with safety experts for the evaluation of work stress. The instrument developed by using these factors had validity, unidimensionality and reliability. The response rate was 81.3%. It is observed that existence the factors responsible for work stress among all the categories of employees in these industries. The multinomial logistic regression model developed is found good in predicting the work stress in manufacturing industries.

KEYWORDS: Work stress, multinomial logistic regression model, manufacturing industries

I. INTRODUCTION

Occupational stress is becoming a major problem in both corporate and social sectors. In industrialized countries, there have been quite dramatic changes in the conditions of work, during the last decade due to the economic, social and technical development. As a consequence the people today at work are exposed to high quantitative and qualitative demands at the work place. In multinational companies, lean production, and down sizing has raised stress level of employees [1-3]. The national institute of occupational safety and health (NIOSH-USA) defines stress as “the harmful physical and emotional responses that occur when the requirements of the job does not match with the capabilities, resources of the workers.” [4,5]

The cost associated with work place stress indicates an international trend among industrialized countries. A recent report says that work related ailments due to work related stress are likely to cost India's exchequer around ₹72000 Crores in 2009-15 [6]. Though India is a fast developing country it is yet to create facilities to mitigate the adverse effects of work stress. The study of work stress in the member states of European Union (EU) points out that an average of 22% of the working Europeans experience work stress [7-9].

It is noted that work stress occurs among the employees at the context of work and at the content of work [10, 11]. The potential stressors for these hazards in the context of work are organizational culture and function, role in the organization, career development, decision latitude and control, interpersonal relationship at work, work-home interface and change [11-13].

Studies on the employees perceptions and descriptions of their organizations, suggest three distinct aspects of organizational function and culture: organization as a task environment, as a problem solving environment and as a development environment [11, 14]. The available evidence suggests that

the organization is perceived to be poor in respect to these environments, will likely to be associated with higher stress [11, 14, 15].

Another major source of stress is associated with person's role at work. A great deal of research is done on role ambiguity and role conflict. It is found that role conflict and role ambiguity are instrumental in developing physiological disorders and says that the above factors can also lead to organizational dysfunction and decreased productivity [16-18]. Lack of expected career growth is one of main sources of work stress. The factors connected with this are poor promotion policies, job insecurity and poor pay in the organization [19-21].

Decision latitude and control are important aspects of work stress. These show the extent which the employees are participating in the decision making process, and also shows the freedom given to the employees for choosing their work [22-24].

The number of research works points out the need of good relationship with superiors, support from the superiors and support from the colleagues at work for the elimination of work related stress hazards [25-28].

Many literature points out the work related stress hazards due to work-family conflict. It is found that that work-family conflict is a form of inter role conflict, in which the role pressures from the work family domains are mutually non compatible in same respect [29-31]. Change is one of the most commonly found stressor at the context of work [11]. It is observed that changes in the modern work environment as result of technological advances, organizational restructuring and various redesign options can elevate the work stress [32, 33].

Like context of work, content of work are also leads to work stress. These factors arise due to improper design of the task, work load and workplace, and work schedule [11,13,34,35]. There are several aspect of job content, which are found hazardous and these include low value of work, low use of skills, repetitive work, uncertainty, lack of opportunity to learn, high attention demand, conflicting demand, insufficient resources [11]. The research work shows that, work related stress hazards arise due to meaning less task and lack of variety etc.... It is also noted that most stressful type of work are those which have excessive demand and pressures that do not match with the workers knowledge and abilities [11,36-38].

The studies on the effect of work stress among men and women working groups in USA and found that due to high psychological work demands like excessive work load and time pressures leads to work stress and cause depression and anxiety in young working adults [11,39-41].

Two major factors responsible for work stress due to the improper work schedule are shift work and long working hours. The studies conducted in Italy among the shift workers observed that shift work leads to poor sleep and health related problems [42-46]. Studies conducted among white collar workers in Sweden, points out that work stress is associated with men subjected to long working hours (75 hours/week) and it is shown that this leads to wide range of ill health in men and women [42-46].

Several models have been proposed to explain the causes of work related stress [47-49]. Frankenhaeuser have described a model where stress is defined in terms of imbalance between the perceived demands from the environment and individuals perceived resources to meet those demands [50]. This imbalance can be caused by quantitative overload (A very high work pace, too much work to do etc...) or qualitative overload (too much responsibility, problems too complex to solve, conflicts etc...).

A well known model describing work stress or strain is the demand control model proposed by Karasek and Theorell and developed and expanded by others. According to this model, the combination of high demands and lack of control and influence (low job discretion) over the work situation causes high work strain [51, 52].

Johannas Siergrist proposed a new model for stress at the work called the effort-reward imbalance model. According to this model, lack of adequate reward in response to the individual's achievement efforts is considered to contribute to high stress levels and elevated health risks. Reward could be obtained in terms of economic benefits, such as higher income [53,54]. Recently multinomial logistic regression models are used to establish relationship between psychosocial work stress factors like work content, work load and social support and job burn out [55-57].

II. SUBJECTS

Total number of subjects selected for this study is 830 and the resulted sample consists of Engineers (75 Nos.), Supervisors (110 Nos.) and workers (675 Nos.). Participants selected for this study consists of both male and female employees of age between 25 to 55 and had sufficient educational background for their job. All employees are permanent and working in shifts in rotation and each shift consists of 8 hour duration per day. However the majority of the employees, in these industries were males and number of woman participants is about 10% of the male participants. All the industries are large scale and profit making for the last five years and located at different districts of Kerala, India. .

III. METHODS

From the literature review and with the consultation of safety experts seven factors were identified for the evaluation of work stress in the absence of well defined factors for the evaluation of work stress in Kerala, India. They are demand, control, manager support, peer support relationship, role and change. The final draft of the questionnaire had 35 items with seven subscales. All the questions were likert type with five fixed alternatives (always, often, sometimes, rarely, never). In addition to this 10 demographic questions are also included in the questionnaire. This questionnaire was refined and validated further by means of confirmatory factor analysis (CFA)[58]. This resulted in removal of five items from the questionnaire. The number of retained items in the questionnaire were demand (7 items), control (4 items), manager support (4 items), peer support (4 items), relationship (4 items), role (5 items) and change (2 items). The values of Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and Cronbach alpha shows that the refined scale has good validity and unidimensionality in addition to reliability [58-60]. The analysis was performed by using the software AMOS-7 [61]. The filled up schedules are then carefully edited for completeness, consistency and accuracy. The overall response rate was 81.3%.

On the basis of data so collected, the influence of factors on works stress analysis is performed using one-way ANOVA. Multinomial logistic regression modelling was done further to find the association of factors responsible for work stress in manufacturing industries.

IV. RESULTS

4.1. Correlation Matrix

A correlation analysis between the variables /factors so identified was performed and the result of the analysis is given in the Table-1. It is noted that all the correlations were positive, but no significant correlation was found between the variable /factors (<0.5). Therefore the variable selected for the study can be treated as independent variables for the purpose of research. The correlation analysis were carried out by means of SPSS-15.

Table 1 : Correlation between the factors

Variables/Factors	Demand	Control	Manager support	Peer support	Relationship	Role	Change
Demand	1	0.354	0.249	0.240	0.310	0.214	0.196
Control	0.354	1	0.279	0.227	0.310	0.168	0.251
Manager support	0.249	0.279	1	0.426	0.319	0.313	0.357
Peer support	0.240	0.227	0.426	1	0.498	0.313	0.461
Relationship	0.310	0.310	0.319	0.498	1	0.440	0.474
Role	0.214	0.168	0.313	0.313	0.440	1	0.353
Change	0.196	0.251	0.357	0.461	0.474	0.353	1

4.2 .Influence of factors on different categories of employees

The influences of these factors are analyzed among different categories of employees by means of one-way ANOVA. The result of the test is given in the Table -2 .The tests are conducted for 0.5 level significance.

Table 2 : Mean Score of Factors

Variables/Factors		Designation			F-value	P-value
		Engineer	Supervisor	Worker		
Demand	Mean	25.72	26.08	25.61	0.603	0.548
	S.D	3.78	4.36	4.11		
Control	Mean	15.09	13.70	12.32	16.644	< 0.001
	S.D	2.96	3.85	4.35		
Manager support	Mean	14.64	15.19	13.94	5.953	0.003
	S.D	2.96	3.85	4.35		
Peer support	Mean	15.94	16.10	15.34	3.748	0.024
	S.D	1.87	2.98	3.12		
Relation-ship	Mean	16.28	17.01	16.41	2.035	0.131
	S.D	1.98	2.82	3.04		
Role	Mean	22.79	22.93	22.56	1.187	0.306
	S.D	2.03	2.14	2.56		
Change	Mean	6.69	6.77	7.02	1.313	0.270
	S.D	1.73	2.34	2.04		

The mean score of the factors /variables points out that existence of factors responsible for work stress among all the categories of the employees in these industries.

It is noted that , significant difference in the factors , control, manager support, and peer support($p < 0.05$) among different categories of employees To identify which among the categories has significant difference , Tukey's multiple comparison test for each of the factors and the results are given in the Table -3

Table -3. Significant difference between different categories of employees

Factors/Variables	Difference between different designation levels
Control	Engineer and worker Supervisor and worker
Manager support	Supervisor and worker
Peer support	Supervisor and worker

The post- hoc analysis, reveals that considerable difference in the mean score of the factor “control” exists between engineers and worker. Further a noted difference is observed for this factor between supervisor and worker .While analyzing the variables manger support and peer support considerable difference is observed only between supervisors and workers

4.3. Multinomial logistic regression modeling of work stress

Multinomial logistic regression is the extension of (binary) logistic regression [62], when categorical dependent outcome has more than one level. An attempt is made to establish multinomial logistic regression model by developing the relationship with the factors/predictor variables .The factors/predictor variables are demand, control, manager support, peer support, relationship, role and

change. The dependent variables for this study are different designation levels namely, engineers, supervisors and workers in the selected manufacturing industries.

For this modelling the reference category is chosen as worker and the multinomial odds of improvement in work stress due to unit increase in the predictor variables/ factors among engineers over the workers is analyzed by this model [see equation---4.3(1)]. Similarly the multinomial odds of improvement in work stress due to unit increase in the predictor variables among supervisors over the workers is analyzed [see equation----- 4.3.(2)]. This relative improvement in work stress for one group over the reference group is explained by means of odds ratio (OR), which is shown under the column Exp(B) (See Table -5). The odds ratio for the event at 95% confidence level is also evaluated. The model fitting information (Table-4), reveals that the initial log likelihood value obtained for the model with no independent variables (intercept only model) is 1076.51. The final log likelihood value obtained for the model by considering all independent variables is 1003.57. The chi-square value obtained is 72.95. As the p-value obtained is below 0.05, we can conclude that the final model is better than the intercept only model.

Table-4. Model fitting information –different designation levels

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	p-value
Intercept Only	1076.51			
Final	1003.57	72.947	14	<0.01

The associations of predictor variables are found with different designation level of the employees. For this analysis, we choose base category as workers

Table-5 . Parameter estimates

Designation		B	Exp(B)	95% confidence interval for Exp(B)	
				Lower bound	Upper bound
Engineer	Intercept	-4.002			
	Demand	.068	1.070	.869	1.004
	Control	.240	1.272	1.165	1.388
	Manager support	.074	1.076	.983	1.179
	Peers support	.073	1.075	.969	1.193
	Relationship	-.124	.884	.790	.989
	Role	.064	1.066	.938	1.211
	Change	-.215	.807	.693	.939
Supervisor	Intercept	-4.548			
	Demand	.016	1.016	.928	1.043
	Control	.076	1.079	1.019	1.143
	Manager support	.104	1.109	1.028	1.197
	Peer support	.063	1.065	.979	1.160
	Relationship	.011	1.012	.915	1.119
	Role	.046	1.047	.943	1.162
	Change	-.230	.795	.704	.897

The reference category : Worker.

The variable shown in the Table-5, has two parts labelled with out come variable designation .This correspond to two equations shown below

$$\begin{aligned} &\text{Log}(p(\text{designation}=\text{engineers})/p(\text{designation}=\text{workers})) \\ &= -4.002 + .068\text{De} + 0.240\text{Cl} + 0.074\text{Ms} + 0.073\text{Ps} - 0.124\text{Re} + 0.064\text{Rl} - 0.215\text{Ch} \end{aligned} \quad (1)$$

$$\begin{aligned} &\text{Log}(p(\text{designation}=\text{supervisors})/p(\text{designation}=\text{workers})) \\ &= -4.548 + 0.016\text{De} + 0.076\text{Cl} + 0.104\text{Ms} + 0.063\text{Ps} + 0.011\text{Re} + 0.046\text{Rl} - 0.230\text{Ch} \end{aligned} \quad (2)$$

V. DISCUSSION

The main aim of the study is to develop and analyze the factors responsible for work stress among the employees in the public sector manufacturing industries in Kerala, India. Accordingly seven factors were developed and the validity, and unidimensionality of the questionnaire was analyzed by means of CFA and the overall reliability of the questionnaire was found satisfactory (>0.70). Interestingly it is found that the factors responsible for work stress is prominent in different categories of employees namely engineers, supervisors and workers these industries. It is also noted that lack of control among lower categories of employees particularly among workers compared to other categories of employees. The results of many earlier research supports the finding [52].

The multinomial logistic regression models can be interpreted by means of odds ratio as

One unit increase in the variable- demand the multinomial odds of improvement of work stress among engineers over workers is expected to increase by a factor 1.070. Similar trend is obtained on supervisors over worker ($\text{OR} > 1$), while keeping all other variables constant [62].

Similarly one unit increase in variables –control, manager support, and peer support ,the multinomial odds of improvement of work stress expected to increase among engineers over workers by the factors 1.272, 1.076 and 1.075 respectively. Similar trend is noticed for supervisors over workers ($\text{OR} > 1$).

The multinomial odds of improvement in work stress, for the unit change in the variable –relationship is expected to increase among supervisors over workers is expected to increase by a factor of 1.012 and reverse trend is noticed for the same variable among engineers over workers ($\text{OR} < 1$).

One unit increase in the variable – role the multinomial improvement in work stress is expected to increase among engineers over workers by a factor 1.066 and similar trend is noticed for this variable among supervisors over the reference group

The multinomial odds of improvement in work stress for the unit change in the variable – change is expected to decrease among engineers over the workers by a factor 0.807 and similar trend is noticed for this variable among supervisors over workers ($\text{OR} < 1$). This model can be used to predict work stress among the employees in manufacturing industries.

Like any other research, the study also not free from limitations. The present study is limited only to public sector industries in Kerala, India, where majority of employees are males. Therefore it would be inappropriate to draw conclusions about male and female workers based on this result. The conclusion is drawn based on the data obtained by means of self reported measures. A comparative study was not carried out because of lack of literature or study of work stress in the context of Indian public sector industries.

VI. CONCLUSION

Consistent with the literature, the results indicate that existence of factors responsible for work stress among all the categories of the employees working in manufacturing industries in Kerala, India and the instrument developed for the evaluation of work stress by using the variables / factors ,namely demand ,control, manager support, peer support, relationship, role and change had validity, unidimensionality and reliability and the instrument can be effectively used for the evaluation of work stress in different type of industries in addition to manufacturing industries . Low level of job control was noticed among lower designation level particularly among workers than engineers and

supervisors. The multinomial logistic regression models proposed are good in representing work stress in the manufacturing industries.

ACKNOWLEDGEMENT

Discussions with Mr. Jacob Devassy, President, International Efficiency Institute, Kochi – 27, India (Subsidiary of International Safety Institute Incorporated, Toronto, Canada) is gratefully acknowledged.

REFERENCES

- [1].International Labour Organization (ILO) report on work stress, 2005
- [2].TheEconomicTimesdated10May2009,<http://eonomictimes.indiatimes.org>
- [3].European Agency for Safety and Health at Work (EASHAW), European risk observation report, 2005.
- [4].T. Cox, A. Griffiths and E. R. Gonzalez,(2000) “Research on work related stress,” European agency for safety and health at Work, Official publication of European communities, Luxemburg .
- [5].C. J. Mackay, R. Cousins, P.J. Kelly, S. Lee and R. H. McCaig, (2004) “Management standards and work related stress in UK policy background and science,” Work and Stress, Vol.18,No.2, pp 91-112.
- [6].J. Park, “Work stress and job performance, (2007)” Perspective – Statistics Canada .Catalogue No.75-001XE, pp 5-17.
- [7].International Labor Organization (ILO) report on work stress,2005
- [8].European agency for safety and health at work (EASHAW) catalogue,2008
- [9].J. Palmer, C. Cooper and K. Thomas ,(2004) “A model of work stress,” Counseling at work .An HSE Publication,1-4.
- [10]. T. Hicks, M. Caroline, (2006) “A Guide to Managing Work Stress,” Florida , Universal publishers
- [11]. T.Cox , A. Griffiths and E.R. Gonaltz (2000).Research on Work Related Stress. European Agency For Safety and Health at Work, Official publication of European communities, Luxemburg.
- [12]. National Institute of Occupational Safety and Health at Work (NIOSH). The Changing Organization of Work .An NIOSH publication (2002), 1-42.
- [13]. R.Cousins, S.D. Clarke, C. Kelly, P.J. Kelly, R.H McCiag and C.J.Mackay (2004)”Management standards and work stress in UK: Practical development” Work and Stress, Vol.18, pp 113-116.
- [14]. T. Cox and I. Howarth (1990) “ Organizational health, culture and helping ,” Work and Stress ,Vol.4,pp 107-110
- [15]. T. Cox and M. Leiter ,(1992) “ The health of the health care organizations ,” Work and Stress ,Vol.6, pp 219-227.
- [16]. T. W. Colligan and, E.M. Higgins, (2005) “Work place stress etiology and consequence,” Journal of Work place Behavioral Health, Vol.2,No.2,pp 89-100.
- [17]. Y.T. Tang and C.H. Chang (2010) “ Impact of role ambiguity and role conflict on employee creativity,” Vol.4,No.6,pp869-881
- [18]. J. M .Stellman ,(1998) “Encyclopedia of Occupational Health and Safety,Vol.2,ILO publication
- [19]. M. Severke and J. Hellgren,(2002) “The nature of job insecurity ,understanding employment uncertainty on the brink of new millennium ,Applied Psychology, Vol.51,No.1,pp 23-52
- [20]. H. Rehman,(2008) “Occupational stress and functional area of an organization,” International Review of Business Research Papers ,Vol.4, No.4, pp 163-173
- [21]. H. Bosma ,R. Peter, J. Siegrist and M. Marmot (1998) “Two alternative job stress models and risk of coronary heart disease,” American Journal of Public Health ,Vol.88,No.1,pp 68-74
- [22]. A. Aras, G. Horgen, H.H. Bjorset, O.Ro, H.Walls and J.Park, (2001) “Masculosketeal ,Visual and psychosocial stress in VDU operators before and after multidisiplinary ergonomic interventions ;A prospective study-part-II,” Applied Ergonomics, Vol.32,pp 559-571.
- [23]. F. Bond and D. Bunce, (2005) “Final Progress Report for British Occupational Health Research Foundation, pp 1-24.
- [24]. M. Turcotte, and G. Schellenberg,(2005) “Job strain and retirement,” Statistics Canada Catalogue No. 75,pp 13-17
- [25]. C.D. Spielberger, P.R. Vagg and C.F. Wasala ,(2003) “Occupational stress, Job pressures and lack of support” in J.C. quick,L.E. Tetrick (Edn.), Hand Book of Occupational Health Psychology, pp 185-200
- [26]. C. Ben , (2007) “FIT work demand and work supports”, pp 1-3.
- [27]. D.D .Bacquer, E.Pelferne, E. Clays, R. Mark, M. Moreeau, P. deSmet, M. Kornitzer and G.De Backer ,(2005) “ Perceived job stress and incidence of coronary events ;A three year follow up of Belgian job stress project cohort,” American Journal of Epidemiology, Vol.16,No.5,pp 431-441

- [28]. P. Brough and J. Williams, (2007) "Managing Occupational stress in a high risk industry: Measuring the job demands of correctional officers," *Criminal Justice and Behavior*, Vol. 34, No. 4, pp 555-567
- [29]. H. Yang, P. L. Schnall, M. Jauregui, T. C. Su and D. Backer (2006) "Work hours and self reported hypertension among working people in California," *Hypertension*, Vol. 48, pp 744-750.
- [30]. N.W.H. Jansen, I.J. Kant, L.G.P.M. Van Amelsvoort, T.S. Kristensen, G.M.H. Swaen, and G.M.H. Nijhuis, 2006 "Work family conflict as a risk factor for sickness absence," *Occupational Environment Medicine*, Vol. 63, pp 488-494.
- [31]. M.R. Frone, (2000) "Work family conflict and employees disorders, The national co-morbidity survey," *Journal of Applied psychology* Vol. 85, No. 6, pp 888-895.
- [32]. J. Shigemitsu, Y. Mino, T. Tsuda, A. Babazono and M. Aoyama, (1997), "The relationship between job stress and mental health at work," *Industrial Health* Vol. 35, pp 29-35.
- [33]. K. Launis and J. Pihlaja, (2007) "Changes in production concepts emphasize problems in work-related well being," *Safety science*, Vol. 45, pp 603-619.
- [34]. T. Cox, (1990) "Stress coping and health psychology," *Work and Stress*, Vol. 4, pp 107-110
- [35]. S. Leka, A. Griffiths and T. Cox, (2003) "Work organization and work stress," *Protecting Workers Health Series*, No. 3, pp 1-32.
- [36]. J. Handy, (1995) "Rethinking stress: Seeing collective," In T. Newton, J. Handy, S. Fineman (Eds.), *Managing stress at work place – Emotion and Power at Work*, New Delhi, Sage
- [37]. *Work-Life and Human Capital Solutions* Available online 10th December 2010
- [38]. T. Chandola, A. Britton, E. Brunner, H. Hemingway, M. Mallic, Mennakumri, E. Badrick, M. Kivimaki, and M. Marmot, "Work stress and coronary heart disease: What are Mechanisms?," *European Journal of Advance Access*, pp 1-9
- [39]. D.A. Whetten and K.S. Cameron, (2007) "Developing and Managing Skills," Seventh edition, Person Education
- [40]. K. Wilkins and M.P. Beaudet, (1998) "Work Stress and Health Reports"- Statistics-Canada Catalogue No. 82-00310(3), pp 47-62.
- [41]. Health and Safety Executive –UK, "How to Tackle Work Related Stress: A Guide for Employers on Making the Management Standards at Work," 2009 pp 1-105.
- [42]. M. Shields, (2006) "Stress and depression in the employed population" *Statistics Canada Health Reports* Vol. 17, No. 4, pp 11-29.
- [43]. Health and Safety Executive –UK, "Essentials of Health and Safety at Work," HSE books, 2006, pp 1-105
- [44]. H. Takeyama, T. Itani, N. Tachi, O. Sakamura, K. Murata, T. Inoue, Takanishi, H. Susumara, and S. Niwa, (2005) "Effect of shift schedules on fatigue and physiological functions among fire fighters during night duty," *Ergonomics*, vol. 48, No. 1, pp 1-11.
- [45]. T. Hirose, (2005) "An Occupational health physicians report in the improvement in sleeping conditions of night shift workers," *Industrial Health*, Vol. 43, pp 58-62.
- [46]. N. Yang, C.C. Chen, J. Choi and Y. Zou, (2000) "Sources work family conflict, A SINO-US comparison of the effects of work and family demands," *Academy of Management Journal* Vol. 43, No. 1, pp 113-123.
- [47]. Y.H. Chan, (2005), "Biostatics 308 Structural equation modeling," *Singapore Medical Journal*, Vol. 46, no. 2, pp 675-679.
- [48]. G. Forgarty, M. Machin, M. Albion, L. Sutherland, G.I. Lalor, and S. Revitt, (1999) "Predicting occupational strain and job satisfaction: The role of stress, coping, and negative affectivity," *Journal of Vocational Behavior*, Vol. 54, pp 429-452.
- [49]. S.Y. Lee, (2007) "Structural equation modeling: A Bayesian approach," *Wiley Series*
- [50]. T. M. Frankenhaeser, (1986) "A Psychobiological frame work for research on human stress and coping," in M.H. Appley, R. Thurnbell, (Ed.), *Dynamics of Stress: Physiological, Psychological, and Social Perspectives*, Plenum, New York, pp 101-116.
- [51]. R. Karasek and T. Theorell, (1990) "Healthy work, stress productivity and the reconstruction of working life", USA Basic books.
- [52]. R. Karasek (1998) "Demand- Control/support model," A social, emotional and physiological approach to stress risk and active behavior development, in J.M. Stellman, (Ed.), *Encyclopedia of Occupational Health and Safety*, Geneva, ILO, Vol. 2, pp 62-69.
- [53]. J. Siegrist, D. Starke, T. Chandola, I. Godwin, M. Marmot, I. Weidhammer and R. Peter, (2004) "Measurement of Effort Reward Imbalance at Work: European Comparisons," *Social Science and Medicine*, Vol. 58, No 8, pp 1483-1499.
- [54]. J. Siegrist, (1996) "Adverse health effects of high effort low reward conditions at work," *Journal of Occupational Health Psychology*, Vol. 1, pp 27-43.

- [55]. R.M. Leontaridi and R.M. Ward, (2002) "Work related stress, quitting intentions and absenteeism," IZA discussion paper No. 493 Germany, pp 1-23.
- [56]. A. Kouvonen, M. Kivimäki, M. Virtanen, J. Pentti, and J. Vahtera, (2005) "Work stress smoking status and smoking intensity: An Observed study of 46190 employees," Journal of Epidemiology and Community Health, Vol. 59, pp 63-69.
- [57]. K.M. Lindblom, S.J. Linton, C. Fedli, and I.L. Bryngelsson (2006) "Burn out in the working population: Relations to the Psychosocial work factors," International Journal of Behavior Medicine, Vol.13, No.1, pp51-59.
- [58]. G. Ntemeyer, W.O. Bearden, and S. Sharma, (2003) "Scaling Procedures –Issues and Applications," New Delhi, Sage.
- [59]. L. J. Cronbach, and P. E. Meehl, (1955) "Construct validity in psychological tests," Psychological Bulletin, Vol. 52, pp 281-302.
- [60]. S. L. Ahire, D.Y. Golhar and M. A. Waller (1996) "Development and validation of TQM implementation construct," Decision Science, Vol. 27, No.1, pp 23-56.
- [61]. J. L. Arbuckle (2006) "AMOS 7.0" AMOS Users Guide. Chicago, IL: Smallwaters corporation.
- [62]. Hosmer. D.W and Lemeshow. S (2000) "Applied logistic regression," second edition, John Wiley & sons, Canada, pp1-369.

Authors

K. Satheesh Kumar is working as Assistant Professor (Senior Grade), in the Department of Mechanical Engineering, Federal Institute of Science and Technology, Kerala, India. He has got about 20 years of experience in teaching and published research papers in national and international journals.



G. Madhu is working as Professor and Head in the Division of Safety and Fire Engineering, Cochin University of Science and Technology, Kerala, India. He has got about 25 years of experience in teaching and published research papers in national and international journals. He has guided many M.Tech and Ph.D students.



AUDIO DENOISING USING WAVELET TRANSFORM

B. JaiShankar¹ and K. Duraiswamy²

¹Department of Electronics and Communication Engineering, KPR Institute of Engineering and Technology, Coimbatore, India.

²Dean, K.S.Rangasamy College of Technology, Tiruchengode, India.

ABSTRACT

Noises present in communication channels are disturbing and the recovery of the original signals from the path without any noise is very difficult task. This is achieved by denoising techniques that remove noises from a digital signal. Many denoising technique have been proposed for the removal of noises from the digital audio signals. But the effectiveness of those techniques is less. In this paper, an audio denoising technique based on wavelet transformation is proposed. Denoising is performed in the transformation domain and the improvement in denoising is achieved by a process of grouping closer blocks. The technique exposes each and every finest details contributed by the set of blocks and also it protects the vital features of every individual block. The blocks are filtered and replaced in their original positions. The grouped blocks overlap each other and thus for every element a much different estimation is obtained. A technique based on this denoising strategy and its efficient implementation is presented in full detail. The implementation results reveal that the proposed technique achieves a state-of-the-art denoising performance in terms of both signal-to-noise ratio and audible quality.

KEYWORDS: Wavelet Transformation, Block Matching, Grouping, Denoising, Reconstruction

I. INTRODUCTION

Signal processing applications always disturbed by noise and it seems to be a major problem. A nonessential signal gets superimposed over an undisturbed signal. If the regularity of noise lessens, then the method for denoising [12] gets more sophisticated. When a signal pass through equipments and connecting wires it naturally gets added with a noise. Therefore it results in signal contamination. Once a signal is polluted, it is essentially difficult to remove it without altering the original signal. Hence, the basic task in signal processing [15] is denoising of signals. Humming noise from audio equipments and background environment noise, both serves as the major root cause for pollution in audio signals. The objective of audio denoising is attenuating the noise, while recovering the underlying signals. It is accessible in many applications such as music and speech restoration etc

Previous methods, such as Gaussian filters and anisotropic diffusion, denoise the value of a signal based on the observed values neighbouring points. Various authors proposed many global and multiscale denoising approaches [15] in order to overcome this locality property. From the beginning of wavelet transforms in signal processing, it is noticed that the wavelet thresholding focuses a attention in removing noise from signals and images. To remove the wavelet coefficients smaller than a given amplitude and to transform the data back into the original domain, the method has to decompose the noisy data into an orthogonal wavelet basis. A nonlinear thresholding estimator can compute in an orthogonal basis such as Fourier or cosine.

In denoising of the audio signals, the denoised signal obtained after performing wavelet transformation is not totally free from noise, some residue of noise left or some other kinds of noise gets introduced by the transformation that is present in the output signal. Several techniques have

been proposed so far for the removal of noise from an audio signal, yet, the efficiency remains an issue as well as they have some drawbacks in general. In this article, we propose an audio signal denoising technique based on an improved block matching technique in transformation domain. The transformation can achieve a clear representation of the input signal so that the noise can be removed well by reconstruction of the signal. A biorthogonal 1.5 wavelet transform is used for the transformation process. A multi dimensional signal vector is generated from the transformed signal vector and the original vector signal is reconstructed by applying inverse transform. The signal to noise ratio (SNR) is comparatively higher than SNR level of the noisy input signal thus increasing the quality of the signal.

II. LITERATURE REVIEW

Michael T. Johnson et al. [6] have demonstrated the application of the Bionic Wavelet Transform (BWT), an adaptive wavelet transform derived from a non-linear auditory model of the cochlea, to the task of speech signal enhancement. Results measured objectively by Signal-to-Noise ratio and Segmental SNR and subjectively by Mean Opinion Score were given for additive white Gaussian noise as well as four different types of realistic noise environments. Enhancement is accomplished through the use of thresholding on the adapted BWT coefficients and the results were compared to a variety of speech enhancement techniques, including Ephraim Malah filtering, iterative Wiener filtering, and spectral subtraction, as well as to wavelet denoising based on perceptually scaled wavelet packet transform decomposition. Overall results indicated that SNR and SSNR improvements for the proposed approach were comparable to those of the Ephraim Malah filter, with BWT enhancement giving the best results of all methods for the noisiest conditions. Subjective measurements using MOS surveys across a variety of 0 db SNR noise conditions indicate that the quality enhancement competitive with but still lower than results for Ephraim Malah filtering and iterative Wiener filtering, but higher than the perceptually scaled wavelet method.

Mohammed Bahoura and Jean Rouat [7] have proposed a new speech enhancement method based on time and scale adaptation of wavelet thresholds. The time dependency was introduced by approximating the Teager energy of the wavelet coefficients, while the scale dependency was introduced by extending the principle of level dependent threshold to wavelet packet thresholding. The technique does not require an explicit estimation of the noise level or of the a priori knowledge of the SNR, as was usually needed in most of the popular enhancement methods. Performance of the proposed method was evaluated on the recorded speech in real conditions (plane, sawmill, tank, subway, babble, car, exhibition hall, restaurant, street, airport, and train station) and artificially added noise. MEL-scale decomposition based on wavelet packets was also compared to the common wavelet packet scale. Comparison in terms of signal-to-noise ratio (SNR) was reported for time adaptation and time-scale adaptation of the wavelet coefficients thresholds. Visual inspection of spectrograms and listening experiments were also used to support the results. Hidden Markov Models speech recognition experiments were conducted on the AURORA-2 database and showed that the proposed method improved the speech recognition rates for low SNRs.

Ching-Ta and Hsiao-Chuan Wang [8] have proposed a method based on critical-band decomposition which converts a noisy signal into wavelet coefficients (WCs), and enhanced the WCs by subtracting a threshold from noisy WCs in each subband. The threshold of each subband is adapted according to the segmental SNR (SegSNR) and the noise masking threshold. Thus residual noise could be efficiently suppressed for a speech-dominated frame. In a noise-dominated frame, the background noise could be almost removed by adjusting the wavelet coefficient threshold (WCT) according to the SegSNR. Speech distortion could be reduced by decreasing the WCT in speech-dominated subbands. The proposed method could effectively enhance noisy speech which was infected by colored-noise. Its performance was better than other wavelet-based speech enhancement methods in their experiments.

Marián Képesia and Luis Weruaga [11] have proposed new method for time-frequency analysis of speech signals. The analysis basis of the proposed Short-Time Fan-Chirp Transform (FChT) was defined univocally by the analysis window length and by the frequency variation rate, that parameter being predicted from the last computed spectral segments. Comparative results between the proposed Short-Time FChT and popular time-frequency techniques reveal an improvement in spectral and

time–frequency representation. Since the signal can be synthesized from its FChT, the proposed method was suitable for filtering purposes.

Nanshan Li et al. [13] have proposed an audio denoising algorithm on the basis of adaptive wavelet soft-threshold which is based on the gain factor of linear filter system in the wavelet domain and the wavelet coefficients teager energy operator in order to progress the effect of the content-based songs retrieval system. Their algorithm integrated the gain factor of linear filter system and nonlinear energy operator with conventional wavelet soft-threshold function. Experiments demonstrated that their algorithm had important outcome in inhibiting Gaussian white noise and pink noise in audio samples and enhancing the accuracy of the songs retrieval system.

III. PROPOSED METHOD

In this section, the proposed audio denoising technique for the removal of unwanted noises from any audio signal is explained. It is considered that the audio signal is polluted by Additive White Gaussian Noise (AWGN) and the polluted signal is subjected to the removal of noise using the proposed denoising technique. The processes performed in the proposed denoising technique are explained in detail as follows (i) Transformation of the noisy signal to wavelet domain, (ii) Generation of a set of closer blocks (iii) Generating a multidimensional array (iv) reconstructing the denoised audio signal. In the proposed work, initially, the noisy audio signal is subjected to Wavelet Transformation.

Wavelet transformation produces a few significant coefficients for the signals with discontinuities.

Audio signals are smooth with a few discontinuities. Hence, wavelet transformation has better capability for representing these signals when compared to other transformations. Once the signal is transformed to wavelet domain, set of closer blocks are synthesized from it.

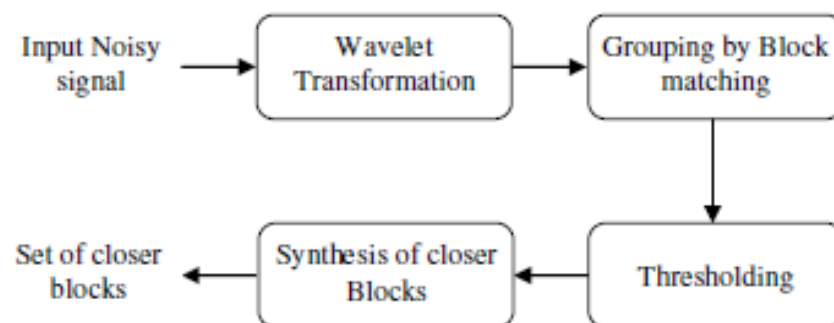


Figure1. Generation of a Set of Closer blocks

3.1. Synthesizing Closer Blocks from the Noisy Signal

The Bior 1.5 wavelet transformation is applied to the input noisy signal and a transformed signal is obtained as output through this process. Designing biorthogonal wavelets allows more level of choices than orthogonal wavelets. One additional level of choice is the possibility to generate symmetric wavelet functions. For the process of transformation, the vector noisy audio signal is reshaped into a matrix of size same as that of the transformation coefficient matrix. The noisy audio signal is transformed to biorthogonal wavelet transformation domain and it is represented as a vector signal that eases the operation.

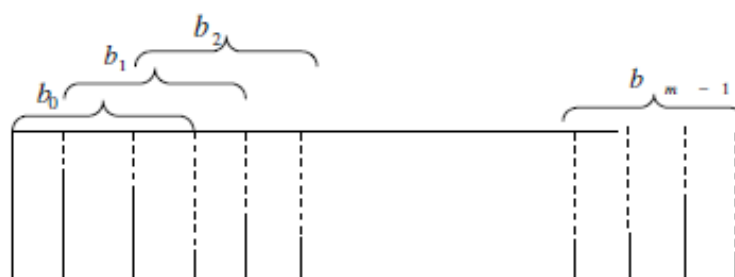


Figure2. Block Representation of the noisy input signal

The process is then followed by the calculation of the L2 norm distance for each block generated in the vector signal. The transformed initial block is kept as the reference block. The distance between

the reference block and all the other grouped blocks are calculated. Similarly, the process is repeated in the same fashion for all the blocks with the consideration of every block as a reference block.

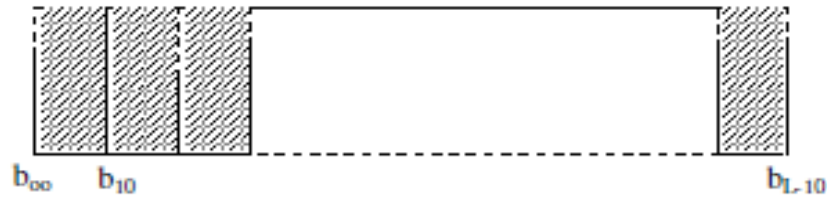


Figure3.Generation of multidimensional vector

The obtained temporary vector signal is then transformed using Haar's transformation. The process of reshaping is carried out to perform the transformation process. Haar's Transformation deals with a 2-point mean and difference operation. Every element in the transformed vector block is compared with a threshold value such that if the element's value is less than the particular element is replaced with '0' and replaced back in the temporary block. If the element's value of the transformed block is greater than the threshold then the value of the transformed element is not changed. The process is repeated for all the elements in the reference block and their respective grouped set of closer blocks. The reconstruction of the signal is shown in figure 4.

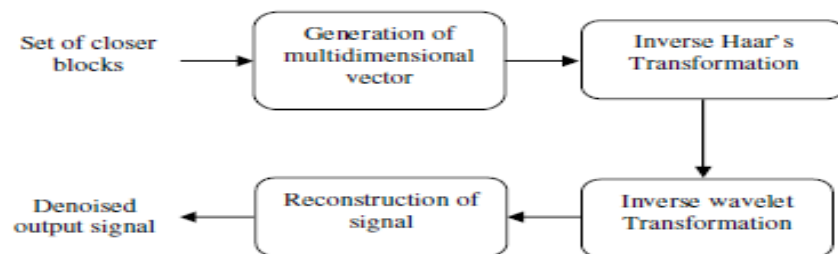


Figure4.Reconstruction of the audio signal

IV. IMPLEMENTATION RESULTS

For testing the performance of the proposed technique a dog barking signal is taken as input. The input signal is then contaminated with AWGN noise for the testing purpose. The input noisy signal with respect to its length $n = 12000$ is represented in Figure 5(a). The AWGN was generated and added to the input signal. The SNR of the noisy audio signal was 7.77 dB at a noise level σ of 0.047. A linear combination of the generated noise and the original signal is used as the primary input for the block matching technique. Figure 5(b) shows the input signal corrupted by white noise. The denoised speech signal achieved through the block matching technique is shown in Figure 5(c). The SNR value of the denoised signal with block matching technique was 12.85 dB.

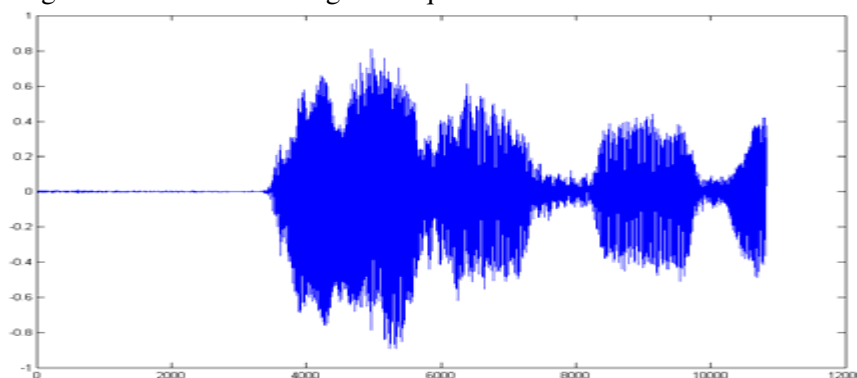


Figure5 (a).Original audio signal

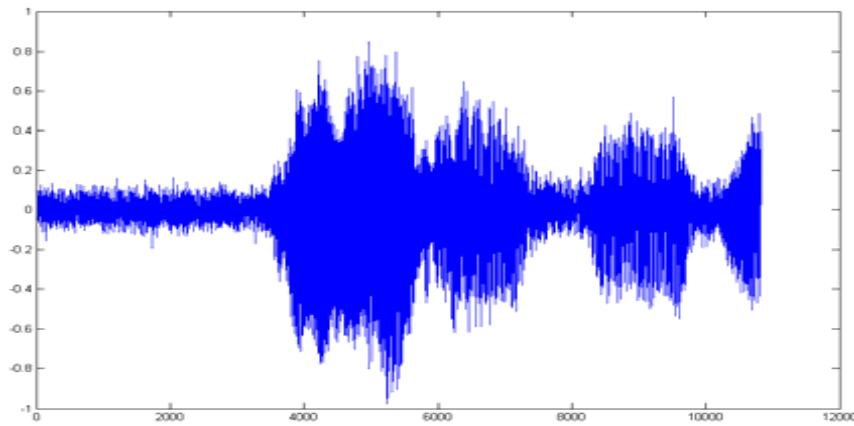


Figure5 (b).Original audio signal with noise

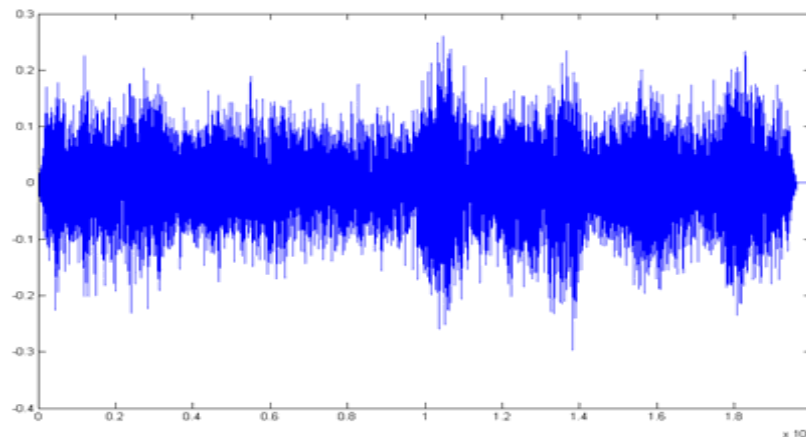


Figure5(c).Denoised audio signal

V. CONCLUSIONS

This paper presented an audio denoising technique based on block matching technique. The technique was based on the denoising strategy and its efficient implementation was presented in full detail. The implementation results have revealed that the process of block matching has achieved a state-of-the-art denoising performance in terms of both peak signal-to-noise ratio and subjective improvement in the audible quality of the audio signal. Grouping of the similar blocks improved the efficient operation of the technique. The blocks were filtered and replaced in their original positions from where they were detached. The grouped blocks were overlapping each other and thus for every element a much different estimation was obtained that were combined to remove noise from the input signal. The reduction in the noise level interprets that the technique has protected the vital unique features of each individual block even when the finest details were contributed by grouped blocks. In addition the technique can be modified for various other audio signals as well as for other problems that can be benefit from highly linear signal representations.

REFERENCES

- [1] Qiang Fu and Eric A. Wan, 2003. "Perceptual Wavelet Adaptive Denoising of Speech", In: *Proc. European Conf. on Speech Commun. and Technology*, pp: 577-580.
- [2] Alyson K. Fletcher, Vivek K Goyal and Kannan Ramchandran, 2003. "Iterative Projective Wavelet Methods for Denoising", *Proc. Wavelets X, part of the 2003 SPIE Int. Symp. on Optical Science & Technology*, Vol. 5207, pp: 9-15, San Diego, CA August.
- [3] Claudia Schremmer, Thomas Haenselmann and Florian Bomers, 2001. "A Wavelet Based Audio Denoiser", In *Proc. IEEE International Conference on Multimedia and Expo (ICME'2001)*, pp: 145-148.
- [4] Sylvain Durand and Jacques Froment, 2001. "Artifact Free Signal Denoising With Wavelets", *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*

- (ICASSP '01), Vol.6, pp: 3685-3688.
- [5] Imola K. Fodor and Chandrika Kamath, 2003. "Denoising Through Wavelet Shrinkage: An Empirical Study", *Journal of Electronic Imaging*, Vol. 12, pp: 151-160.
 - [6] Michael T. Johnson, Xiaolong Yuan and Yao Ren, 2007. "Speech Signal Enhancement through Adaptive Wavelet Thresholding", *Speech Communications*, Vol. 49, No. 2, pp: 123-133.
 - [7] Mohammed Bahoura and Jean Rouat, 2006. "Wavelet speech enhancement based on time-scale adaptation", *Speech Communication*, Vol. 48, No. 12, pp: 1620-1637.
 - [8] Ching-Ta and Hsiao-Chuan Wang, 2003. "Enhancement of single channel speech based on masking property and wavelet transform", *Speech Communication*, Vol. 41, No 2-3, pp: 409-427.
 - [9] Ching-Ta and Hsiao-Chuan Wang, 2007. "Speech enhancement using hybrid gain factor in critical-band-wavelet-packet transform", *Digital Signal Processing*, Vol. 17, No. 1, pp: 172-188.
 - [10] Erik Visser, Manabu Otsuka and Te-Won Lee, 2003. "A spatio-temporal speech enhancement scheme for robust speech recognition in noisy environments", *Speech Communication*, Vol. 41, No. 2-3, pp: 393-407.
 - [11] Marián Képesia and Luis Weruaga, 2006. "Adaptive chirp-based time-frequency analysis of speech signals", *Speech Communication*, Vol. 48, No. 5, pp: 474-492.
 - [12] Claudia Schremmer, Thomas Haenselmann and Florian Bomers, 2001. "A Wavelet Based Audio Denoiser," *Proceedings of IEEE International Conference on Multimedia and Expo*, pp: 145-148, Tokyo, Japan.
 - [13] Nanshan Li and Mingquan Zhou, 2008. "Audio Denoising Algorithm Based on Adaptive Wavelet Soft-Threshold of Gain Factor and Teager Energy Operator," *Proceedings of IEEE International Conference on Computer Science and Software Engineering*, Vol. 1, pp: 787-790.
 - [14] Enqing Dong and Xiaoxiang Pu, 2008. "Speech denoising based on perceptual weighting filter," *Proceedings of 9th IEE International Conference on Signal Processing*, pp: 705-708, October 26-29, Beijing.
 - [15] Amit Singer, Yoel Shkolnisky, and Boaz Nadler, 2009. "Diffusion Interpretation of Nonlocal Neighborhood Filters for Signal Denoising", *SIAM J. Imaging Sciences*, Vol. 2, No. 1, pp: 118-139.
 - [16] Guoshen Yu, Stephane Mallat and Emmanuel Bacry, 2007. "Audio Signal Denoising with Complex Wavelets and Adaptive Block Attenuation," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 3, pp: 869-872, April 15-20, Honolulu, HI.
 - [17] Guoshen Yu, Stéphane Mallat and Emmanuel Bacry, 2008. "Audio Denoising by Time-Frequency Block Thresholding", *IEEE Transactions On Signal Processing*, Vol. 56, No. 5, pp: 1830-1839.
 - [18] D. L. Donoho and I. M. Johnstone, 1994. "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, Vol. 81, pp: 425-455.
 - [19] D. L. Donoho, 1995. "De-noising by soft-thresholding," *IEEE Trans. Inform. Th.*, Vol. 41, pp: 613-627.
 - [20] R. R. Coifman and D. L. Donoho, 1995. "Translation-invariant de-noising, in Wavelets and Statistics", A. Antoniadis and G. Oppenheim, eds., *Springer Lecture Notes in Statistics* 103, pp: 125-150, Springer-Verlag, New York.
 - [21] Jonathan Berger and Charles Nichols, "Brahms at the piano," *Leonardo Mus. Journal*, vol. 4, pp. 23-30, 1994.
 - [22] David L. Donoho, "Nonlinear Wavelet Methods for Recovering Signals, Images, and Densities from Indirect and Noisy Data," <http://wwwstat.stanford.edu/~donoho/Reports/>, 1993.
 - [23] Ingrid Daubechies, *Ten Lectures on Wavelets*, vol. 61, SIAM. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992, ISBN 0-89871-274-2.
 - [24] St'ephaneMallat, *AWavelet Tour of Signal Processing*, Academic Press, San Diego, CA, USA, 1998, ISBN 0-12-466605-1.
 - [25] M. Lang, H. Guo, J.E. Odegard, and C.S. Burrus, "Nonlinear processing of a shift invariant DWT for noise reduction," *SPIE, Mathematical Imaging: Wavelet Applications for Dual Use*, April 1995.
 - [26] Curtis Roads, *The Computer Music Tutorial*, The MIT Press, 1996.
 - [27] Claudia Schremmer, Thomas Haenselmann, and Florian B'omers, "Wavelets in Real-Time Digital Audio Processing: A Software For Understanding Wavelets in Applied Computer Science," in *Workshop on Signal Processing Applications (WoSPA)*, <http://www.sprc.qut.edu.au/wospa2000/>, December 2000, Signal Processing Research Center (SPRC) and IEEE.
 - [28] Claudia Schremmer, "The Wavelet Tool," <http://www.informatik.uni-mannheim.de/~cschremm/>

- wavelets/WaveletAudioTool/, 2000.
- [29] Manojit Roy, V. Ravi Kumar, B.D. Kulkarni, John Sanderson, Martin Rhodes, and Michel van der Stappen, "Simple denoising algorithm using wavelet transform," *AIChE Journal*, vol. 45, no. 11, pp. 2461–2466, 1999.
- [30] Florian B"omers, "Wavelets in Real-Time Digital Audio Processing: Analysis and Sample Implementations," M.S. thesis, Universit"at Mannheim, <http://www.bome.com/personal/thesis.pdf>, 2000.

AUTHORS PROFILE

B. Jai Shankar received B.E. degree in Electronics and Communication Engineering from Government College of Engineering, Salem and M.E. degree in Applied Electronics from Kongu engineering College, Erode. He worked in K.S.R College of Engineering, Tiruchengode for three years. Currently he is working as lecturer in Kumaraguru College of Technology, Coimbatore since 2008. His research interest includes Digital Signal Processing, Image Processing and Wavelets.



K. Duraiswamy received his B.E. degree in Electrical and Electronics Engineering from P.S.G. College of Technology, Coimbatore in 1965 and M.Sc.(Engg) from P.S.G. College of Technology, Coimbatore in 1968 and Ph.D. from Anna University in 1986. From 1965 to 1966 he was in Electricity Board. From 1968 to 1970 he was working in ACCET, Karaikudi. From 1970 to 1983, he was working in Government College of Engineering Salem. From 1983 to 1995, he was with Government College of Technology, Coimbatore as Professor. From 1995 to 2005 he was working as Principal at K.S. Rangasamy College of Technology, Tiruchengode and presently he is serving as DEAN of K.S. Rangasamy College of Technology, Tiruchengode, India. Dr. K. Duraiswamy is interested in Digital Image Processing, Computer Architecture and Compiler Design. He received 7 years Long Service Gold Medal for NCC. He is a life member in ISTE, Senior member in IEEE and a member of CSI.



HYBRID ACTIVE POWER FILTER USING FUZZY DIVIDING FREQUENCY CONTROL METHOD

SaiRam.I¹, Bindu.V² and K.K. Vasishta Kumar³

¹Associate Professor

²Assistant Professor and ³Assistant Professor

Department of Electrical and Electronics Engineering,
Dhanekula Institute of Engineering & Technology, Vijayawada, India

ABSTRACT

This paper deals with a hybrid active power filter with injection circuit (IHAPF). It shows great promise in reducing harmonics and improving the power factor with a relatively low capacity active power filter. This paper concluded that the stability of the IHAPF based on detection supply current is superior to that of others. To minimize the capacity of IHAPF, an adaptive fuzzy dividing frequency-control method is proposed by analyzing the bode diagram, which consists of two control units: a generalized integrator control unit and fuzzy adjustor unit. The generalized integrator is used for dividing frequency integral control, while fuzzy arithmetic is used for adjusting proportional-integral coefficients timely. And the control method is generally useful and applicable to any other active filters. Compared to other IHAPF control methods, the adaptive fuzzy dividing frequency control shows the advantages of shorter response time and higher control precision. It is implemented in an IHAPF with a 100-kVA APF installed in a copper mill in Northern China. The simulation and experimental results show that the new control method is not only easy to be calculated and implemented, but also very effective in reducing harmonics.

KEYWORDS: Hybrid Active Power Filter (HAPF), Fuzzy-control method, Integral controller.

I. INTRODUCTION

Harmonic filters are used to eliminate the harmonic distortion caused by nonlinear loads. Specifically, harmonic filters are designed to attenuate or in some filters eliminate the potentially dangerous effects of harmonic currents active within the power distribution system. Filters can be designed to trap these currents and, through the use of a series of capacitors, coils, and resistors, shunt them to ground. A filter may contain several of these elements, each designed to compensate a particular frequency or an array of frequencies.

A passive filter is composed of only passive elements such as inductors, capacitors and resistors thus not requiring any operational amplifiers. Passive filters are inexpensive compared with most other mitigating devices. Its structure may be either of the series or parallel type. The structure chosen for implementation depends on the type of harmonic source present. Internally, they cause the harmonic current to resonate at its frequency. Through this approach, the harmonic currents are attenuated in the LC circuits tuned to the harmonic orders requiring filtering. This prevents the severe harmonic currents traveling upstream to the power source causing increased widespread problems.

An active filter is implemented when orders of harmonic currents are varying. One case evident of demanding varying harmonics from the power system are variable speed drives. Its structure may be either of the series or parallel type. The structure chosen for implementation depends on the type of

harmonic sources present in the power system and the effects that different filter solutions would cause to the overall system performance. Active filters use active components such as IGBT-transistors to inject negative harmonics into the network effectively replacing a portion of the distorted current wave coming from the load. This is achieved by producing harmonic components of equal amplitude but opposite phase shift, which cancel the harmonic components of the non-linear loads.

Hybrid filters combine an active filter and a passive filter. Its structure may be either of the series or parallel type. The passive filter carries out basic filtering (5th order, for example) and the active filter, through precise control, covers higher harmonics.

II. CLASSIFICATION OF ACTIVE FILTERS ACCORDING TO POWER CIRCUIT, CONFIGURATIONS AND CONNECTIONS

a) Shunt active filters:

Shunt active filters are by far the most widely accepted and dominant filter of choice in most industrial processes. The active filter is connected in parallel at the PCC and is fed from the main power circuit. The objective of the shunt active filter is to supply opposing harmonic current to the nonlinear load effectively resulting in a net harmonic current. This means that the supply signals remain purely fundamental.

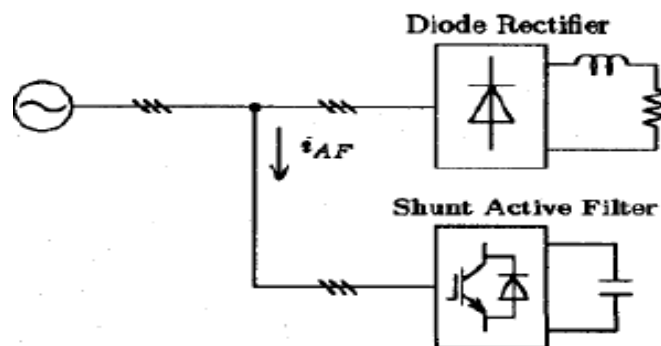


Figure 1 Shunt active filter used alone

Shunt filters also have the additional benefit of contributing to reactive power compensation and balancing of three-phase currents. Since the active filter is connected in parallel to the PCC, only the compensation current plus a small amount of active fundamental current is carried in the unit. For an increased range of power ratings, several shunt active filters can be combined together to withstand higher currents.

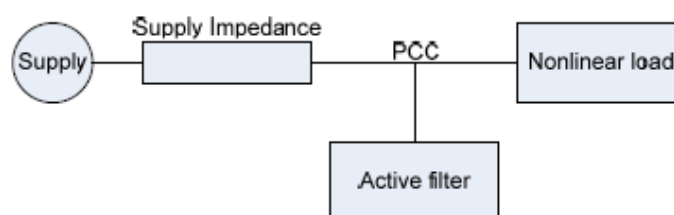


Figure.2 Shunt active filter network configuration

b) Series active filters:

The objective of the series active filter is to maintain a pure sinusoidal voltage waveform across the load. This is achieved by producing a PWM voltage waveform which is added or subtracted against the supply voltage waveform. The choice of power circuit used in most cases is the voltage-fed PWM inverter without a current minor loop. The active filter acts as a voltage source and thus it is often a preferred solution of harmonic producing loads such as large capacity diode rectifiers with capacitive loads. In general, series active filters are less commonly used against the shunt design.

Unlike the shunt filter which carries mainly compensation current, the series circuit has to handle high load currents. This causes an increased rating of the filter suitable to carry the increased current. Series filters offer the main advantage over the shunt configuration of achieving ac voltage regulation by eliminating voltage-waveform harmonics. This means the load contains a pure sinusoidal waveform.

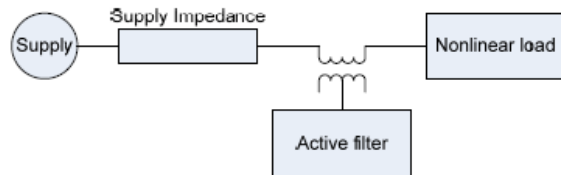


Figure.3 Series active filter configuration

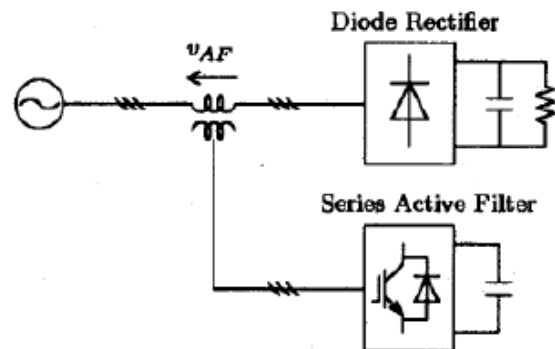


Figure.4 Series active filter used alone

c) *Combination of both shunt and series active filters:*

The diagram shown in figure below shows the combination of both parallel and series active filters. This system combines both the benefits of the shunt and series and is often used to achieve the demanding power system requirements. The control of active filters can be complex. A combination of the two provides an even greater complexity. The higher cost involved in a more complex design has shown a reduced demand for the combined structure. As a result of the increased cost and complexity, this combination has received less attention than other configurations. Flexible AC transmission systems, commonly abbreviated as FACTS regularly make use of the arrangement.

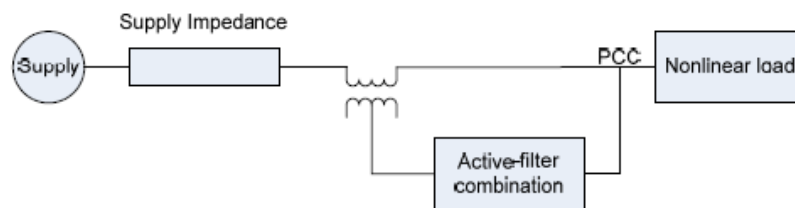


Figure.5 Combination of shunt and series active filters

d) *Combination of series active and shunt passive filters :*

The combination of the active parallel and active series filters in was seen to be very complex in control yielding a high cost. One method of reducing these problems was to replace the parallel active filter with a passive structure. The series active filter, which constitutes high impedance for high-frequency harmonics, is accompanied by a parallel passive filter to provide a path for the harmonic currents of the load. This combination represented by figure below, permits an improvement over the characteristics of plain series active filters and the extension of their capabilities to include current-harmonic reduction and voltage- harmonic elimination. Passive filters are often easier and simple to implement and do not require any control circuit. This, this deserves to be most beneficial.

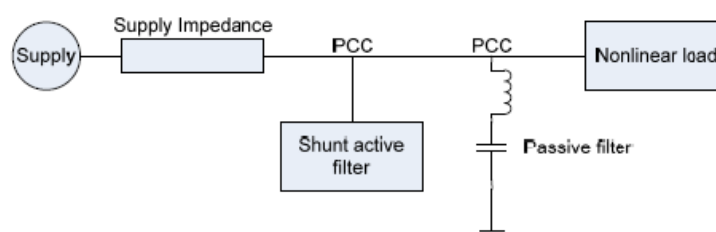


Figure.6 Combination of series active and shunt passive filters

e) *Combination of shunt active and passive filter:*

As mentioned, shunt active filters are best suitable to compensate for lower order harmonics thus only requiring low power rating which serves most economical. This configuration makes use of a passive

filter which serves to compensate for the high order load current harmonics. This combination, represented by figure (shunt active filter network configuration) presents this important configuration. Combinations such as this can be designed to compensate for higher powers without excessive costs for high-power switching. The major disadvantage of this configuration is the fact that passive filters can only be tuned for a specific predefined harmonic and thus cannot be easily changed for loads which have varying harmonics.

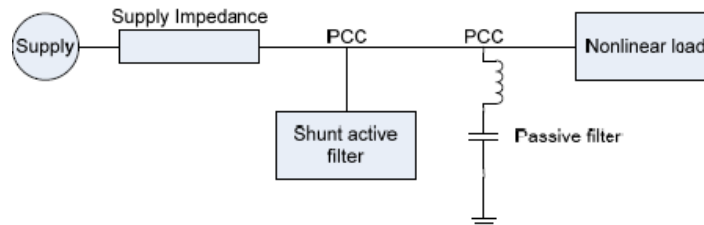


Figure.7 Combination of shunt active and passive filter

f) Active filter in series with shunt passive filter:

The combination of an active filter in series with a shunt passive filter is considered a significant design configuration for medium and high voltage applications. The passive filter is designed to reduce the voltage stress applied to the switches in the active filter. This design is in its infancy of development however, further research is still needed to assess the effectiveness of the configuration.

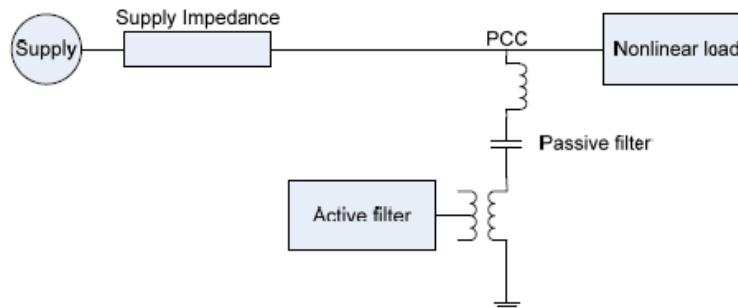


Figure.8 Active filter in series with shunt passive filter

III. SYSTEM CONFIGURATION AND CONTROL STRATEGY

Topology of the Novel HAPF:

The parallel HAPF has the advantages of easy installation and maintenance and can also be made just by transformation on the PPF installed in the grid. Fig.9 shows a PHAPF that is in use now. To reduce the power of APFs, a PPF has been designed for some certain orders of harmonics. As in Fig.9, and make up a PPF to compensate the second, fifth, and seventh harmonic current, while the APF is just used to improve the performance of PPF and get rid of the resonance that may occur.

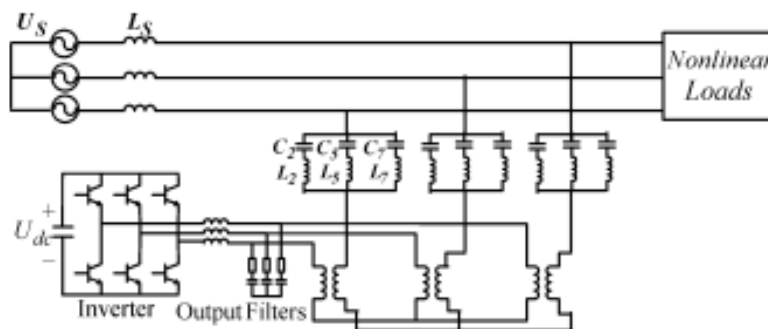


Figure 9.Topology of the Shunt Hybrid APF

So the power of the filter can be reduced sharply, usually one-tenth of the power of the nonlinear load, which enables the APF to be used in a high-power occasion.

HAPF is expected to compensate for reactive power as well as damp harmonics in the grid, and all of the reactive power current will go through APF. To further decrease the power of APF, a novel configuration of the hybrid APF is proposed as shown in Fig. 10 and tune at the fundamental frequency, and then compose the injection branch with the APF, shunted to the fundamental resonance circuit, is directly connected in series with a matching transformer. Therefore, the novel HAPF (IHAPF) is formed. The PPF sustains the main grid voltage and compensates for the constant reactive power, while the fundamental resonance circuit only sustains the harmonic voltage, which greatly reduces the APF power and minimizes the voltage rating of the semiconductor switching device. So it is effective to be used in the 6-kV/10-kV medium-voltage grid.

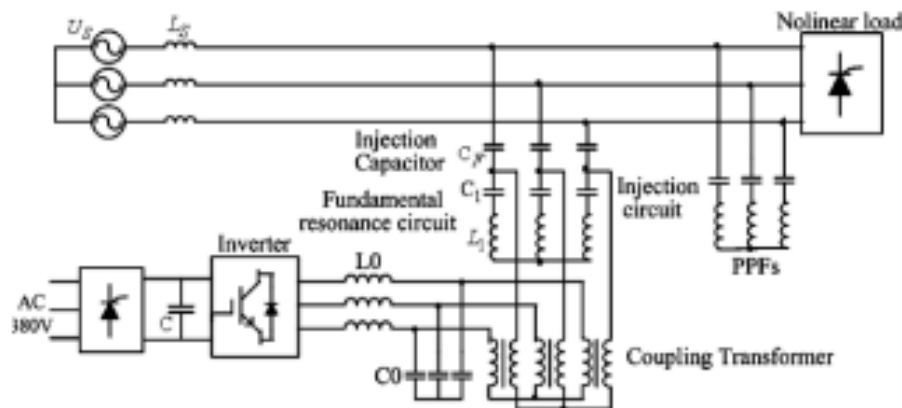


Figure.10 Topology of the Novel HAPF

IV. ADAPTIVE FUZZY DIVIDING FREQUENCY-CONTROL METHOD

The conventional linear feedback controller (PI controller, state feedback control, etc.) is utilized to improve the dynamic Response and/or to increase the stability margin of the closed loop system.

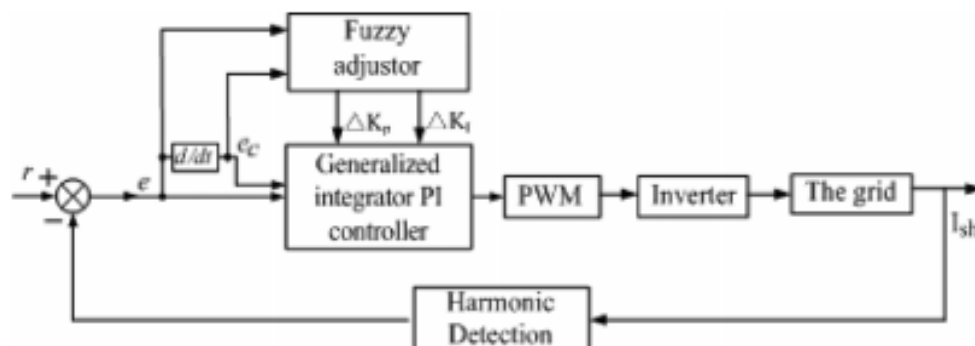


Figure.11 Configuration of the adaptive fuzzy dividing frequency controller

However, these controllers may present a poor steady-state error for the harmonic reference signal. An adaptive fuzzy dividing frequency control method is presented in Fig. 11, which consists of two control units: 1) a generalized integrator control unit and 2) a fuzzy adjustor unit.

The generalized integrator, which can ignore the influence of magnitude and phase, is used for dividing frequency integral control, while fuzzy arithmetic is used to timely adjust the PI coefficients. Since the purpose of the control scheme is to receive a minimum steady-state error, the harmonic reference signal is set to zero. First, supply harmonic current is detected. Then, the expectation control signal of the inverter is revealed by the adaptive fuzzy dividing frequency controller.

The stability of the system is achieved by a proportional controller, and the perfect dynamic state is received by the generalized integral controller. The fuzzy adjustor is set to adjust the parameters of proportional control and generalized integral control. Therefore, the proposed harmonic current tracking controller can decrease the tracking error of the harmonic compensation current, and have better dynamic response and robustness.

V. SIMULATION RESULTS

Simulation of a 10-kV system has been carried out with software MATLAB shown in the figure 12. The system parameters are listed in Table.1.

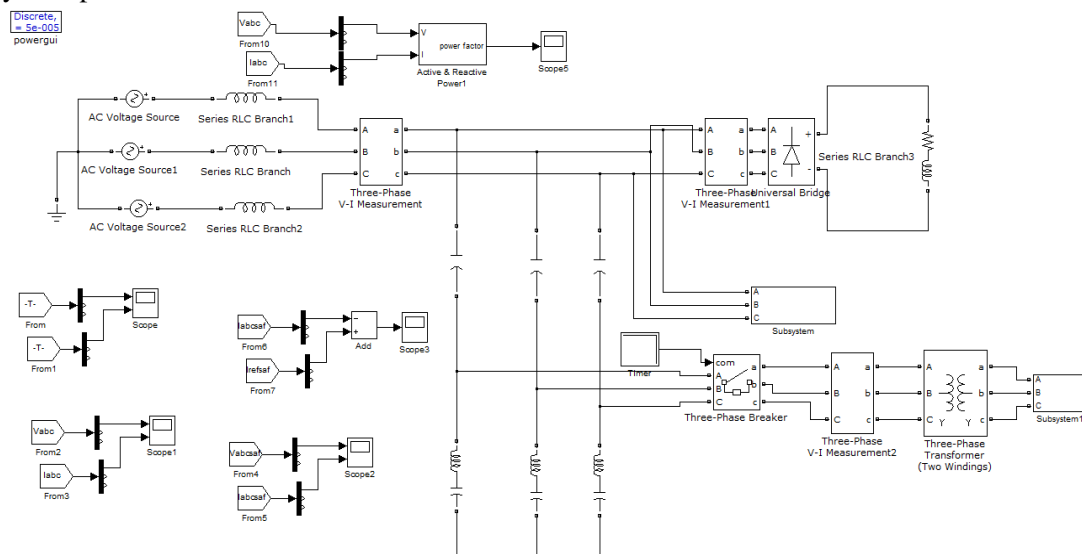


Figure 12: Matlab Simulation Circuit with adaptive fuzzy controller

Table.1: System Parameters

	L/mH	C/ μ F	Q
Out put filter	0.2	60	
11 th turned filter	1.77	49.75	50
13 th turned filter	1.37	44.76	50
6 th turned filter	14.75	$C_F: 19.65, C_I: 690$	

The PPFs are turned at the 11th and 13th, respectively. The injection circuit is turned at the 6th. In this simulation, ideal harmonic current sources are applied. The dc-side voltage is 535 V. Simulation results with the conventional PI controller and the proposed current controller are shown in Fig.13.

I_L , I_s , I_F , I_{apf} , and the error represent the load current, supply current, current through the injection capacitor, current through APF, and error of compensation. When the proposed controller is used, the error can be reduced to 10 A in 0.06 s. It is observed that compared to the conventional PI controller and generalized integral controller, the proposed controller has better dynamic performance.

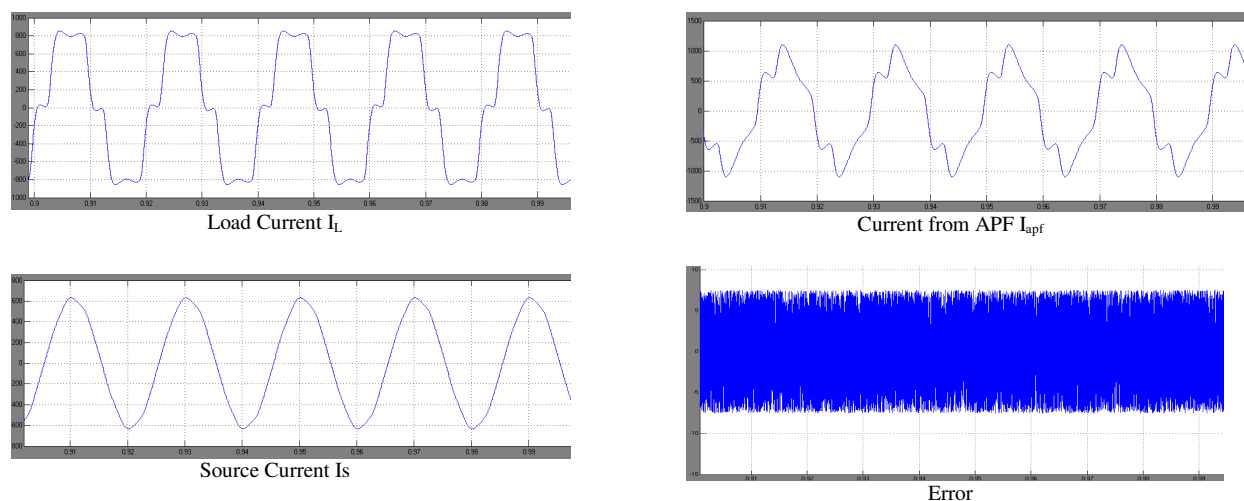


Figure 13 Simulation results of steady state performance with adaptive fuzzy controller

Fig.13 shows the steady-state performance of the IHAPF when fuzzy controller is used. From Fig. 13, it can be seen that after IHAPF with the conventional PI controller runs, the current total harmonic distortion reduces to 3.8% from 21.8%, and the power factor increases to 0.95 from 0.55.

When the conventional generalized integral controller is used, the current THD reduces to 3.3% from 21.8%, while after the IHAPF with the proposed PI controller runs; the current THD reduces to 1.8% from 21.8%. So it can be observed that the proposed current controller exhibits much better performance than the conventional PI controller and the conventional generalized integral controller. Table.2 shows Comparison of supply current THD and Power factor.

Table 2: Comparison without and with IHAPF

	THD	Power Factor
Without IHAPF	21.5%	0.69
With IHAPF	1.9%	0.94

VI. CONCLUSION

A novel hybrid APF with injection circuit was proposed. Its principle and control method was discussed. The proposed adaptive fuzzy-dividing frequency control can decrease the tracking error and increase dynamic response and robustness. The control method is also useful and applicable to any other active filters. It is implemented in an IHAPF with a 100-kVA APF system in a copper mill in Northern China, and demonstrates good performance for harmonic elimination. Simulation and application results proved the feasibility and validity of the IHAPF and the proposed control method.

REFERENCES

- [1] A.Luo, Z.Shuai, W.Zhu, C.Tu "Development of Hybrid Active Power Filter Based on the Adaptive Fuzzy Dividing Frequency-Control Method", IEEE Trans. Power Del, Volume 24, No.1, 424-432, January 2009.
- [2] N. Mohan, H.A. Peterson, W.F. Long, G.R. Dreifuerst, and J.J. Vithayathil, "Active filters for AC harmonic suppression," presented at the IEEE Power Eng. Soc. Winter meeting, 1977.
- [3] F. Peng, H. Akagi, and A. Nabae, "A new approach to harmonic compensation in power system-a combined system of shunt passive and series active filters," IEEE Trans. Ind. Appl., vol.26, no.6, pp.983-990, Nov. 1990.
- [4] C. Madtharad and S. Premrudeepreechacharn, "Active power filter for three-phase four-wire electric systems using neural networks," Elect. Power Syst. Res., vol.60, no. 2, pp.179-192, Apr.2002.
- [5] H. Fujita and H. Akagi, "a practical approach to harmonic compensation in power system-series connection of passive and active filters," IEEE trans Ind. Appl., vol.27, no.6, pp.1020-1025, Nov.1991.
- [6] S. Kim and P.N. Enjeti, "A new hybrid active power filter (APF) topology," IEEE trans. Power Electronics, vol.17, no. 1, pp. 48-54, Jan. 2002.
- [7] A. Nakajima, K. Oku, J. Nishidai, T. Shiraishi, Y.Ogihara, K. Mizuki, and M. Kumazava, "Development of active filter with series resonant circuit," in Proc 19th IEEE Annu. Power Electronics Specialists Conf. Rec., Apr. 11-14 1988, vol. 2, pp. 1168-1173.

Authors Biography:

I. Sai Ram is currently working as Associate Professor in EEE department, Dhanekula Institute of Engineering & Technology, Vijayawada. His research areas include Power Systems, Electrical Machines and Control Systems.



V. Bindu is currently working as Assistant Professor in EEE department, Dhanekula Institute of Engineering & Technology, Vijayawada. Her research areas include Power Systems, Power Quality, power Electronics and Control systems



K. K. Vasishta Kumar is currently working as Assistant Professor in EEE department, Dhanekula Institute of Engineering & Technology, Vijayawada. His research areas include Power Systems, Power Quality and Electrical Machines.



MINIMUM LINING COST OF TRAPEZOIDAL ROUND CORNERED SECTION OF CANAL

Syed Zafar Syed Muzaffar¹, S. L. Atmapoojya², D.K. Agarwal³

¹Assistant Professor (Civil Engg), Anjuman College of Engg. & Technology, Nagpur, India

²Principal, S.B Jain Institute of Technology & Management Nagpur, India

³Ex Dean, Faculty of Engg. & Tech., RSTM & Research University Nagpur, India

ABSTRACT

Economics and Environmental pressure require efficient use of water for irrigation purpose. To satisfy the Economics constrains in canal construction the total cost of section should be minimum which include cost of lining for minimizing the total cost of lined canal trapezoidal shape the Lagrangian multiplier technique is used. The total cost of lining is the sum of cost of lining on sides on curves and in the base. Radius of rounding can be provided in two ways 1) unbound radius case i.e. the radius of rounding remain constant and it predecided 2) Bound radius case i.e. the radius rounding depend on depth of flow. These two cases are only for design purpose of canal section. It is seen that cost of lining is less in case of bound radius case.

KEYWORDS: Unbound radius, Bound radius, side lining, Base lining.

I. INTRODUCTION

Trapezoidal round cornered section is more efficient than sharp cornered section. Attempts have been made by researches to derive the optimal sharp cornered section. The paper deals with minimize the total cost of lining canal system which is provide to convey water from a source to required point. It is found that about 45% to 50% water is lost due to seepage from canal system during its journey from head works to the field. The seepage also enhances water logging of the adjacent area to the canal which causes in reduction of crop production. Hence it required how to control loss of water due to seepage. Then lining of canal system is one of the majors which overcome this problem. Lining of canal prevent soil erosion due to high velocity which increases cost of canal system. A section of unlined canal system does not remain in trapezoidal shape for longer times there is considerable seepage loss through it. Whereas the lined canal section is provided with a hard surface coating around the perimeter of canals which prevent the seepage loss. Hence canal may be designed by using flow equations. To get economy in canal construction the section should be design in such a way that its total cost of lining should be minimum.

The radius of rounding may be taken as constant and independent of depth of flow, which is termed as unbound radius case. In second case radius of rounding is taken as a function of depth of flow which is termed as bound radius case.

The cost of lining material for the base, curves and sides are considered as different. The curved portion need special type of form work and labour cost is also more, similarly on the sides the performance of work will be different from the base, and hence the cost of lining per unit area is considered as different for base, curve and sides.

Let

C_b = Cost of Lining per unit area for base.

C_c = Cost of Lining per unit area for the curved portion.

C_s = Cost of Lining per unit area for the sides.

$$C = bC_b + 2rz_2C_c + (2y\sqrt{1+z^2} - 2rz_1)C_s \quad \text{-----(1)}$$

Round corned section can be

- Unbound radius section
- Bound radius section

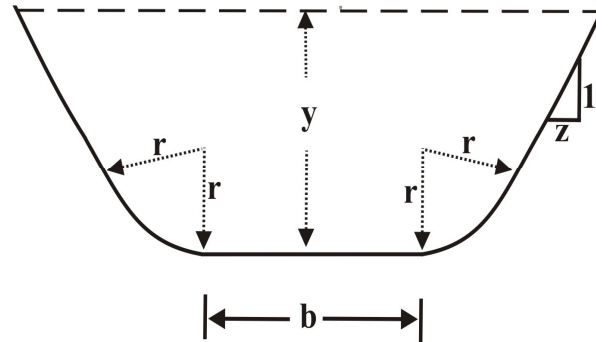


Figure1 Trapezoidal round cornered lined canal section

II. UNBOUND RADIUS CASE

Radius of rounding $r = \text{constant}$

$$A = by + 2ryz_1 + y^2z - 2r^2z_1 + r^2z_2 \quad \text{-----(2)}$$

$$A = y^2 \left(\frac{b}{y} + 2 \frac{r}{y} z_1 + z - 2 \left(\frac{r}{y} \right)^2 z_1 + \left(\frac{r}{y} \right)^2 z_2 \right)$$

$$y^2 = A / \left(\frac{b}{y} + 2 \frac{r}{y} z_1 + z - 2 \left(\frac{r}{y} \right)^2 z_1 + \left(\frac{r}{y} \right)^2 z_2 \right)$$

$$y = \sqrt{A / \left(\frac{b}{y} + 2 \frac{r}{y} z_1 + z - 2 \left(\frac{r}{y} \right)^2 z_1 + \left(\frac{r}{y} \right)^2 z_2 \right)}$$

$$y = [1 \times \sqrt{A}] / \sqrt{\left(\frac{b}{y} + 2 \frac{r}{y} z_1 + z - 2 \left(\frac{r}{y} \right)^2 z_1 + \left(\frac{r}{y} \right)^2 z_2 \right)}$$

$$y = \phi_1 \times \sqrt{(Q/V)} \quad \text{----- (3)}$$

Putting the value of b from equation (2) in equation (1) and we get

$$C = (A/y - 2rz_1 - yz + 2r^2/y z_1 - r^2/y z_2) C_b + 2rz_2C_c + (2y\sqrt{1+z^2} - 2rz_1) C_s \quad \text{----- (4)}$$

For minimum value of C differentiating with respect to y and equating to zero we get

$$b/y = 2 \left[\sqrt{1+z^2} C_s/C_b - z \right] - 2(r/y)z_1 \quad \text{----- (5)}$$

III. BOUND RADIUS CASE

Radius of rounding $r = f(y)$

And let $r = Ky$

Where K = constant its value is less than unity

The flow area

$$A = by + 2ky^2z_1 + y^2z - 2k^2y^2z_1 + k^2y^2z_2 \quad \text{-----(6)}$$

Wetted perimeter

$$P = b + 2y\sqrt{1+z^2} - 2kyz_1 + 2kyz_2 \quad \text{-----(7)}$$

Cost of lining is given by

$$C = bC_b + 2kyz_2C_c + [2y\sqrt{1+z^2} - 2kyz_1] C_s \quad \text{----- (8)}$$

For minimum value of C

Differentiating C with respect to y and equating to zero we get

$$b/y = 2 \left(\sqrt{1+z^2} C_s/C_b - z \right) - 2kz_1 \left(2 + C_c/C_b \right) + 2kz_2 C_c/C_b + 4k^2(z_1 - z_2/2) \quad \text{----- (9)}$$

The value of y can be determined by equation (6) as

$$y = \sqrt{A / \left(\frac{b}{y} + 2kz_1 + z - 2k^2z_1 + k^2z_2 \right)}$$

$$y = [1 \times \sqrt{A}] / \sqrt{\left[2 \left(\sqrt{1+z^2} \right) C_s/C_b - z \right] + z - 2kz_1 \left(1 + C_s/C_b \right) + 2kz_2 C_s/C_b + 2k^2(z_1 - z_2/2)} \quad \text{Or}$$

$$y = \phi_2 \times \sqrt{(Q/V)} \quad \text{-----(10)}$$

Where

$$\phi_2 = 1 / \sqrt{\left[2 \left(\sqrt{1+z^2} \right) C_s/C_b - z \right] + z - 2kz_1 \left(1 + C_s/C_b \right) + 2kz_2 C_s/C_b + 2k^2(z_1 - z_2/2)} \quad \text{Or}$$

$$\phi_2 = f(C_s/C_b, k, z)$$

After obtaining the value of y for a given value of b can be obtained discharge and velocity of flow.

Example:

$$\begin{aligned}
 Q &= 90 \text{ m}^3/\text{Sec} & K &= 0.7 \\
 V &= 1.8 \text{ m/Sec} \\
 Z &= 1.5 \\
 C_b &= \text{Rs. } 200 / \text{Sq M} \\
 C_s &= \text{Rs } 500 / \text{Sq M} \\
 r/y &= 0.7 \\
 C_c &= 225 / \text{Sq M} \\
 Z_1 &= \sqrt{(1+z^2)} - z & \sqrt{(1+z^2)} &= 1.8027 \\
 &= \sqrt{(1+1.5^2)} - 1.5 \\
 Z_1 &= 0.30277 & \text{As } Q/v &= 90/1.8 = 50 \\
 Z_2 &= \tan^{-1}(1/z) \times \pi / 180 \\
 Z_2 &= \tan^{-1}(1/1.5) \times \pi / 180 \\
 Z_2 &= 0.588
 \end{aligned}$$

Unbound Radius Case:

$$\begin{aligned}
 b/y &= 2(\sqrt{(1+z^2)} C_s/C_b - z) - 2 r/y z_1 \\
 &= 2(\sqrt{(1+1.5^2)} 500/200 - 1.5) - 2 \times 0.7 \times 0.30277 \\
 b/y &= 5.59 \\
 \text{The value of factor, } \phi_1 &\text{ for calculated value of } b/y \text{ and given value of } r \text{ and } z \\
 \Phi_1 &= 1 / \sqrt{(b/y + 2z_1 r/y + z - 2 r^2/y^2 z_1 + r^2/y^2 z_2)} \\
 \Phi_1 &= 1 / \sqrt{(5.59 + 2 \times 0.30277 \times 0.7 + 1.5 - 2 \times 0.7^2 \times 0.30277 + 0.7^2 \times 0.588)} \\
 &= 1 / \sqrt{(5.59 + 0.42387 + 1.5 - 0.2967 + 0.288)} \\
 &= 1 / 2.739 \\
 \Phi_1 &= 0.365 \\
 \text{The depth of flow } y & \\
 y &= \Phi_1 \sqrt{(Q/V)} \\
 &= 0.365 \sqrt{(90/1.8)} \\
 &= 2.581 \text{ meter} \quad \text{and} \quad r = r/y \times y \\
 & \quad \quad \quad = 0.7 \times 2.58 \\
 & \quad \quad \quad r = 1.8067 \\
 \text{and } b &= 2.58 \times 5.56 \\
 b &= 14.428 \text{ meter} \\
 \text{Cost of lining } C & \\
 C &= b C_b + 2 r z_2 C_c + (2y \sqrt{(1+z^2)} - 2 r z_1) C_s \\
 &= 14.428 \times 200 + 2 \times 1.8067 \times 0.588 \times 225 + (2 \times 2.581 \times \sqrt{(1+1.5^2)} - 2 \times 1.8067 \times 0.30277) \times 500 \\
 C &= 7469.602 \text{ per meter}
 \end{aligned}$$

Bound Radius Case:

$$\begin{aligned}
 b/y &= 2(\sqrt{(1+z^2)} C_s/C_b - z) - 2kz_1(2 + C_s/C_b) + 2kz_2 C_c/C_b + 4k^2(z_1 - z_2/2) \\
 &= 2(1.8027 \times 500/200 - 1.5) - 2 \times 0.7 \times 0.30277(2 + 500/200) + 2 \times 0.7 \times 0.588 \times 225/200 + 4 \times 0.7^2(0.30277 - 0.588/2) \\
 b/y &= 5.054 \\
 \Phi_2 &= 1 / \sqrt{[2(\sqrt{(1+z^2)} C_s/C_b - z) + z - 2kz_1(1 + C_s/C_b) + 2kz_2 C_c/C_b + 2k^2(z_1 - z_2/2)]} \\
 &= 1 / \sqrt{[2(1.8027 \times 500/200 - 1.5) + 1.5 - 2 \times 0.7 \times 0.30277 \times (1 + 500/200) + 2 \times 0.7 \times 0.588 \times 225/200 + 2 \times 0.7^2(0.30277 - 0.588/2)]} \\
 \Phi_2 &= 0.2949 \\
 y &= \Phi_2 \sqrt{(Q/V)} \\
 &= 0.2949 \times \sqrt{(90/1.8)} \\
 y &= 2.085 \text{ meter} & b &= 10.54 \text{ meter} \\
 C &= b C_b + 2kyz_2 C_c + (2y \sqrt{(1+z^2)} - 2kyz_1) C_s \\
 &= 10.54 \times 200 + 2 \times 0.7 \times 2.085 \times 0.588 \times 225 + (2 \times 2.085 \times 1.8027 - 2 \times 0.7 \times 2.085 \times 0.30277) \times 500
 \end{aligned}$$

$C = \text{Rs } 5810.83 \text{ per meter}$

IV. CONCLUSION

The condition for optimal design of trapezoidal canal is developing by considering the cost of lining. The cost of lining which is the sum of cost of base lining, cost of side lining and cost of curve lining is minimise with respect to variable for given area of flow. The optimal cost obtained by the proposed method by considering radius as bound and unbound. In unbound radius case we take radius as a constant. In bound radius case we take radius as function of depth of flow. It is seen that the cost obtained by bound radius is less than unbound radius.

REFERENCES

- [1] Prabhata K Swamee, Govinda C, Mishra and Bhagur R, Chahur "Minimum Cost Design of Lined Canal Section" Water Resource Management 14 1-12-2000.
- [2] M.Riaz and Z Sen European "Aspects of Design and benefits of Alternative lining Systems" European Water 11/12, 17-27, 2005 @ 2005. EW Publications.
- [3] S.K Garg "A text book of Irrigation Engineering and Hydraulic Structures" member of Indian Water Resource Society.
- [4] Dr. B.C Punmia and Dr. Pandey B.B Lal "A text book of Irrigation and Water Power Engineering".
- [5] IS 10430-1982 "Criteria for Design of Lined Canal and Guide for selection of type of Lining" Bureau of Indian Standards, Manak Bhavan, New Delhi.
- [6] IS 4515-1993 "stone Pitched Lining for Canals- Code of Practice" Bureau of Indian Standards, Manak Bhavan New Delhi.
- [7] V. T. Chow "Open Channel Hydraulics" The McGraw Hill book company New York Nu, 1959.
- [8] P. N. Modi and S. M. Seth "Design of Most Economical Trapezoidal Section of open Channels" Journal of Irrigation and Power, New Delhi, 1968, pp 271-280.
- [9] C. Y. Guo and W. C. Hughes "Optimal Channel Cross Section With Free Board" Journal of Irrigation and Drainage Division ASCE Vol. 110, no.8, 1984, pp 304-314.
- [10] A Das "Optimal Channel Cross Section with composite roughness" Journal of Irrigation and Drainage Division ASCE Vol. 126, no.1, 2000, pp 68-71.
- [11] G. V. Loganathan "Optimal Design of Parabolic Canals" Journal of Irrigation and Drainage Division, ASCE Vol.117 no.5, 1991, pp 716-735.
- [12] P. Monadjemi "General Formulation of Best Hydraulic Channel Section" Journal of Irrigation and Drainage Division ASCE Vol.120, no.1, 1994, pp 27-35.
- [13] V. L. Streeter and E. B. Wylie "Fluid Mechanics" Mc Graw Hills Inc, New York Ny, 1979.
- [14] S. L. Atmapoojya and R. N. Ingle "The Optimal Canal Section With Consideration of Free Board" IE (I) Journal- CV, Vol.83, Feb-2003.

Author

Syed Zafar Syed Muzaffar was born on 08-07-1964. He has completed his B.E. (Civil Engg.) in 1990 from B. N. College of Engg, Pusad Dist. Yavatmal, M. Tech. (Hydraulic Engg.) in 1999 from V.R.C.E., Nagpur & Pursuing Ph.D. Civil Engg. from Nagpur University. He is working as Assistant Professor in Anjuman College of Engineering, Sadar, Nagpur (India)



S.L. Atmapoojya has completed B.E. (CIVIL) from Government College of Engineering, Jabalpur, M.Tech (Hydraulic) from VRCE, Nagpur & Ph D - From RSTM, Nagpur university, Nagpur. He is presently working as Professor in KITS, Ramtek, Nagpur.

Dinesh Kumar Onkarnath Agrawal was born on 15.08.1960. He has completed his B.Sc. from Bhopal University (M.P.) in 1980, M.Tech in 1983 from Dr. Harisingh Gour University, Sagar (M.P.) & Ph.D. in 2007 from RTM Nagpur University Nagpur (M.S.). He is working as Professor in Faculty of Engineering & Technology, RSTM & Research University Nagpur, India.

VOLTAGE CONTROL AND DYNAMIC PERFORMANCE OF POWER TRANSMISSION SYSTEM USING STATCOM AND ITS COMPARISON WITH SVC

Amit Garg¹ and Sanjai Kumar Agarwal²

¹Assistant Engineer, Uttar Haryana Bijli Vitran Nigam Limited, Haryana.

²Professor EEE Deptt., YMCA University of Science & Technology, Faridabad.

ABSTRACT

This paper investigates the effects of Static Synchronous Compensator (STATCOM) on voltage stability of a power system. This paper will discuss and demonstrate how STATCOM has successfully been applied to power system for effectively regulating system voltage. One of the major reasons for installing a STATCOM is to improve dynamic voltage control and thus increase system load ability. This paper presents modeling and simulation of STATCOM in MATLAB/Simulink. In this paper A STATCOM is used to regulate voltage in a power system. When system voltage is low the STATCOM generates reactive power (STATCOM capacitive). When system voltage is high it absorbs reactive power (STATCOM inductive). The STATCOM more effectively enhance the voltage stability and increase transmission capacity in a power system. In this paper comparison is also performed between SVC and STATCOM under fault condition and it is shown that STATCOM have the ability to provide more capacitive power during a fault over SVC. It is also shown that STATCOM exhibits faster response than SVC.

KEYWORDS: Dynamic Performance, FACTS, Matlab/Simulink, Transient Stability.

I. INTRODUCTION

Today's changing electric power systems create a growing need for flexibility, reliability, fast response and accuracy in the fields of electric power generation, transmission, distribution and consumption. Flexible Alternating Current Transmission Systems (FACTS) are new devices emanating from recent innovative technologies that are capable of altering voltage, phase angle and/or impedance at particular points in power systems. Their fast response offers a high potential for power system stability enhancement apart from steady state flow control. Among the FACTS controllers, STATCOM provides fast acting dynamic reactive compensation for voltage support during contingency events which would otherwise depress the voltage for a significant length of time. In emerging electric power systems, increased transactions often lead to the situations where the system no longer remains in secure operating region. The flexible AC transmission system (FACTS) controllers can play an important role in the power system security enhancement. However, due to high capital investment, it is necessary to locate these controllers optimally in the power system. FACTS devices can regulate the active and reactive power control as well as adaptive to voltage-magnitude control simultaneously because of their flexibility and fast. In this paper Control System Block diagram and V-I Characteristics of STATCOM are described. STATCOM is modeled in MATLAB/ Simulink and simulation results are shown and discussed.

II. STATIC SYNCHRONOUS COMPENSATOR (STATCOM)

The Static Synchronous Compensator (STATCOM) is a shunt device of the Flexible AC Transmission Systems (FACTS) family using power electronics to control power flow and improve transient stability on power grids. The STATCOM regulates voltage at its terminal by controlling the amount of reactive power injected into or absorbed from the power system. When system voltage is low, the STATCOM generates reactive power (STATCOM capacitive). When system voltage is high, it absorbs reactive power (STATCOM inductive).

III. CONTROL SYSTEM BLOCK DIAGRAM OF STATCOM

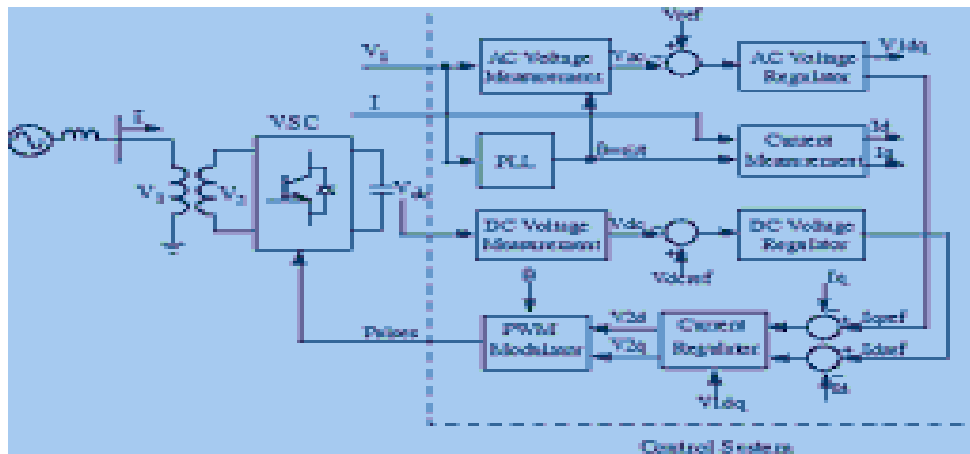


Fig 1: Block Diagram of STATCOM

Control System Block Diagram of STATCOM

The control system consists of:

- A phase-locked loop (PLL) which synchronizes on the positive-sequence component of the three-phase primary voltage V_1 . The output of the PLL (angle $\theta = \omega t$) is used to compute the direct-axis and quadrature-axis components of the AC three-phase voltage and currents (labeled as V_d , V_q or I_d , I_q on the diagram).
- Measurement systems measuring the d and q components of AC positive-sequence voltage and currents to be controlled as well as the DC voltage V_{dc} .
- An outer regulation loop consisting of an AC voltage regulator and a DC voltage regulator. The output of the AC voltage regulator is the reference current I_{qref} for the current regulator (I_q = current in quadrature with voltage which controls reactive power flow). The output of the DC voltage regulator is the reference current I_{dref} for the current regulator (I_d = current in phase with voltage which controls active power flow).
- An inner current regulation loop consisting of a current regulator. The current regulator controls the magnitude and phase of the voltage generated by the PWM converter (V_{2d} V_{2q}) from the I_{dref} and I_{qref} reference currents produced respectively by the DC voltage regulator and the AC voltage regulator (in voltage control mode). The current regulator is assisted by a feed forward type regulator which predicts the V_2 voltage output (V_{2d} V_{2q}) from the V_1 measurement (V_{1d} V_{1q}) and the transformer leakage reactance.

IV. V-I CHARACTERISTICS OF STATCOM

As long as the reactive current stays within the minimum and maximum current values ($-I_{max}$, I_{max}) imposed by the converter rating, the voltage is regulated at the reference voltage V_{ref} . However, a voltage droop is normally used (usually between 1% and 4% at maximum reactive power output), and the V-I characteristic has the slope indicated in the figure. In the voltage regulation mode, the V-I characteristic is described by the following equation:

$$V = V_{ref} + X_s I$$

Where

V : Positive Sequence Voltage (pu)

I : Reactive Current ($I > 0$ indicates an Inductive Current)

X_s : Slope or Droop Reactance

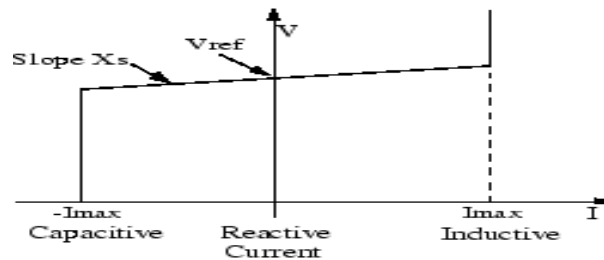


Fig 2: V-I Characteristics of STATCOM

V. MODELLING OF STATCOM/SVC

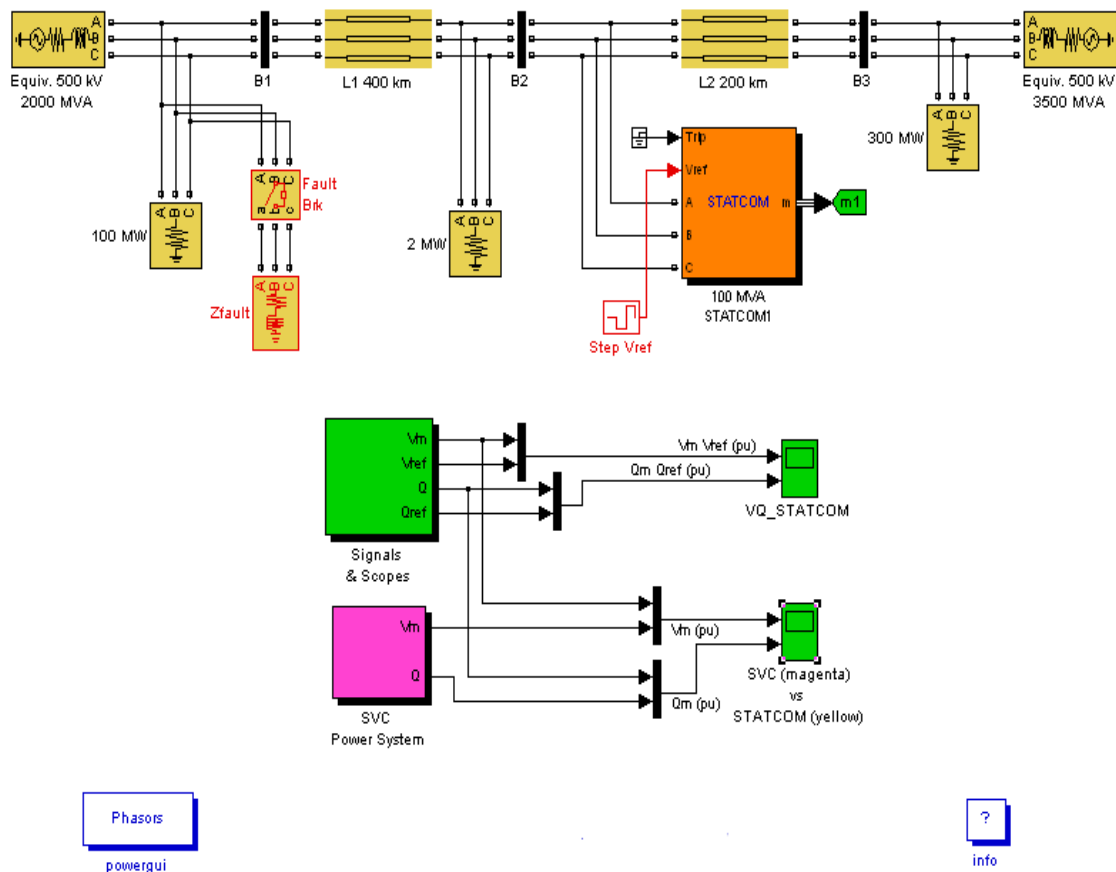


Fig 3: STATCOM / SVC on a 500 KV Transmission Line

The power grid consists of two 500-kV equivalents (respectively 2000 MVA and 3500 MVA) connected by a 600-km transmission line. In this paper STATCOM has a rating of ± 100 MVA. "Droop" parameter should be set to 0.03 and the "Vac Regulator Gains" to 5 (Proportional gain K_p) and 1000 (Integral gain K_i). "Step Vref" block (the Red timer block connected to the "Vref" input of the STATCOM) should be programmed to modify the reference voltage V_{ref} as follows: Initially V_{ref} is set to 1 pu; at $t=0.2$ s, V_{ref} is decreased to 0.97 pu; then at $t=0.4$ s, V_{ref} is increased to 1.03; and finally at 0.6 s, V_{ref} is set back to 1 pu. Here fault breaker at bus B1 will not operate during the simulation.

VI. COMPARISON OF STATCOM AND SVC

In another experiment for comparison of STATCOM model with a SVC model having the same rating (± 100 MVA) SVC is connected to a power grid similar to the power grid on which our STATCOM is connected and disabling the "Step Vref" block by multiplying the time vector by 100 then program the fault breaker by selecting the parameters "Switching of phase A, B and C" and verify that the breaker is programmed.

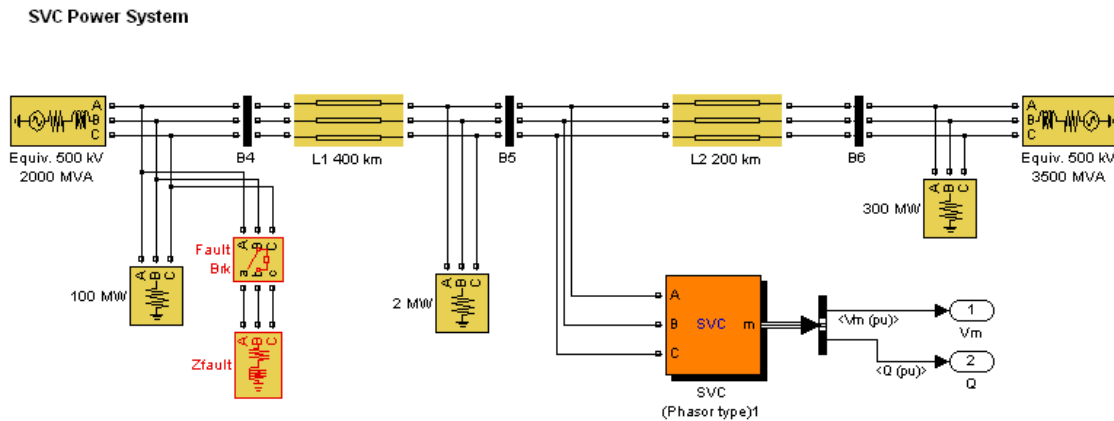


Fig 4: SVC Power System

VII. SIMULATION RESULTS AND DISCUSSIONS

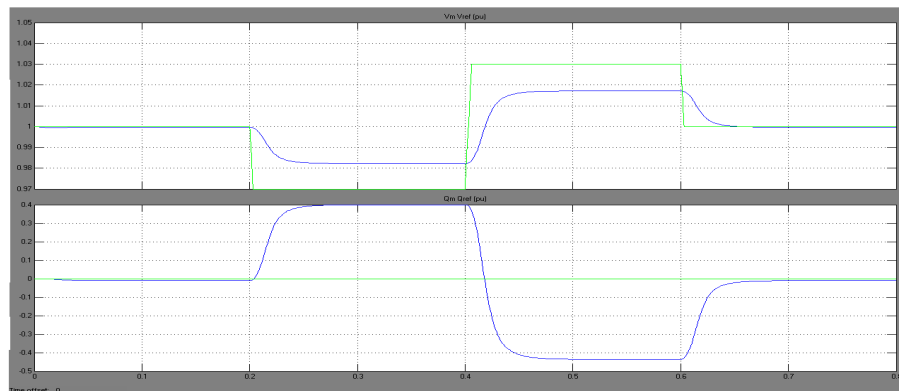


Fig 5: STATCOM Dynamic Response

The first graph displays the V_{ref} signal (green trace) along with the measured positive-sequence voltage V_m at the STATCOM bus (violet trace). The second graph displays the reactive power Q_m (violet trace) absorbed (positive value) or generated (negative value) by the STATCOM. The signal Q_{ref} (green trace) is not relevant to this simulation.

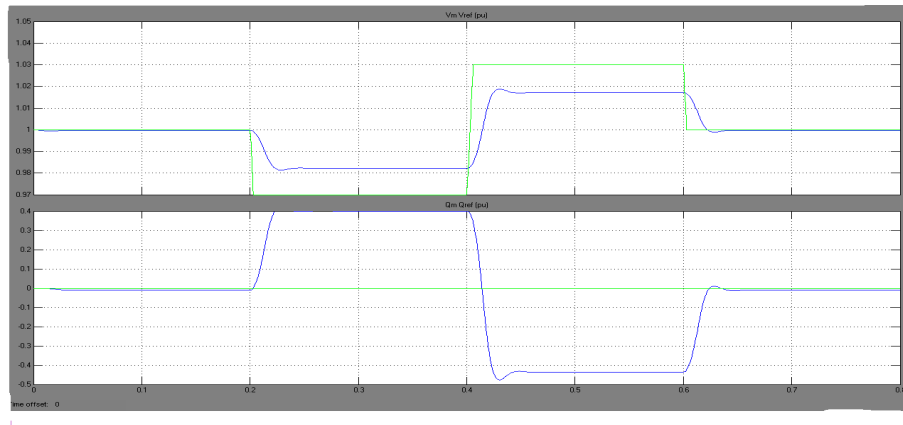


Fig 6: STATCOM Faster Dynamic Response

To check the impact of the regulator gains, the two gains of the Vac Regulator Gains are multiplied by two and simulation is rerun. A much faster response with a small overshoot is observed as shown in Fig. 6.

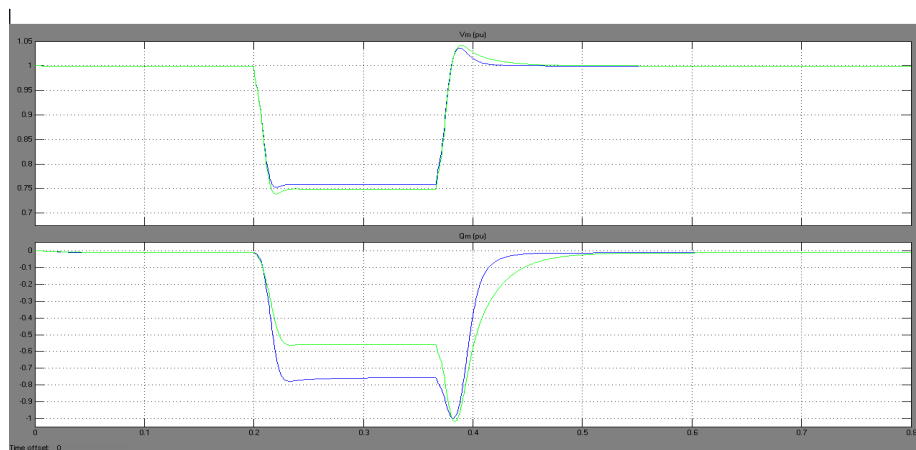


Fig 7: Comparison between Vm and Qm of SVC and STATCOM

For comparison of STATCOM model with a SVC model SVC is connected to a power grid similar to the power grid on which STATCOM is connected and disabling the "Step Vref" block by multiplying the time vector by 100 then program the fault breaker by selecting the parameters "Switching of phase A, B and C". Now simulation is run again at the "SVC vs STATCOM" scope graphs shown in Fig. 7 are observed. The first graph displays the measured voltage Vm on both systems (green trace for the SVC). The second graph displays the measured reactive power Qm generated by the SVC (green trace) and the STATCOM (violet trace). During the 10-cycle fault, a key difference between the SVC and the STATCOM can be observed. The reactive power generated by the SVC is -0.58 pu and the reactive power generated by the STATCOM is -0.79 pu.

VIII. CONCLUSION

Hence it is concluded that the maximum capacitive power generated by a SVC is proportional to the square of the system voltage (constant susceptance) while the maximum capacitive power generated by a STATCOM decreases linearly with voltage decrease (constant current). This ability to provide more capacitive power during a fault is one important advantage of the STATCOM over the SVC. In addition, the STATCOM will normally exhibit a faster response than the SVC because with the voltage-sourced converter, the STATCOM has no delay associated with the thyristor firing (in the order of 4 ms for a SVC). So STATCOM provides fast acting dynamic reactive compensation for voltage support during contingency events which would otherwise depress the voltage for a

significant length of time. So In this paper it is successfully demonstrated that how STATCOM has successfully been applied to power system for effectively regulating system voltage.

REFERENCES

- [1] Nang Sabai, and Thida Win (2008) "Voltage control and dynamic performance of power transmission system using SVC" World Academy of Science, Engineering and Technology 42 Pp. 425-429
- [2] P.Kundur, "Power system stability and control", Mc Graw-Hill, 1994
- [3] D. Murali (2010),"Comparison of FACTS devices for power system stability enhancement" International Journal of Computer Applications (0975 – 8887) Volume 8– No.4, pp. 30-35
- [4] H. Yazdanpanahi , "Application of FACTS devices in transmission expansion to overcome the problems related to delays".
- [5] A.E. Hammad, "Analysis of power system stability enhancement by static var compensator", IEEE PWRS, vol 1, no 4, pp. 222-227.
- [6] Christian Rehtanz April (2009) , "New types of FACTS devices for power system security and efficiency" Pp-1-6.
- [7] Hadi Saadat, "Power system analysis", Mc Graw-Hill companies. Inc,1999.
- [8] Edris Abdel, "Series Compensation Schemes Reducing the Potential of Sub synchronous Resonance, "IEEE Trans. On power systems, vol. 5 No. 1. Feb1990. Pp. 219-226
- [9] Haque M.H (1992).," Maximum transfer capability with in the voltage stability limit of series and shunt compensation scheme for AC transmission systems", Electric Power system research, vol. 24, pp. 227-235
- [10] Hauth R.L., Miske S.A. and Nozari F, (Oct 1982).," The role and benefits of static VAR systems in High Voltage power system applications", IEEE trans on PAS, Vol PAS-101, pp. 3761-3770.

Authors:

Amit Garg: BE(EE) from Punjab Engineering College, Chandigarh. Post Graduation in Instrumentation and Control from Apeejay College of Engg & Technology, Gurgaon is working as Assistant Engineer in Uttar Haryana Bijli Vitran Nigam, Haryana.



Sanjai Kumar Agarwal: B.Tech (EE) from NIT Calicut in 1984. Post Graduation in Control Systems from Delhi College of Engineering in 2000. He has completed his Ph.D.in Electrical Engineering from Jamia Milia Islamia University, New Delhi in 2009.His research interests include Flexible AC Transmission Controllers.

ASSOCIATION RULE MINING ALGORITHMS FOR HIGH DIMENSIONAL DATA – A REVIEW

K.Prasanna¹ and M.Seetha²

¹ JNIAS Research Scholar, Asstt. Prof., Department of CSE, AITS, Rajampet, A.P., India

² Professor in CSE, GNITS, Hyderabad, Hyderabad. India

ABSTRACT

In this paper the association rule mining algorithms are discussed and demonstrated. Particularly, the problems of association rule mining on high dimensional data are investigated and comparison of popular association rules algorithms are presented. The paper mainly focusing on the problem of curse of dimensionality associated with data bases and algorithms. To day there are several efficient algorithms that cope with the task of Association rule mining. Actually, these algorithms are less described in terms of dimensionality. In this paper, we described to day's approaches pointing out the common aspect and differences. A comprehensive experimental study against different UCI data sets are evaluated and presented. It turns out that the run time behavior of the algorithms with regards to numerous dimensionalities, derived rules, and processing time similar to be expected. It was ascertained that as increase in the dimensionality of the databases and with varying support values, the run time performance is proven better and faster in CARPENTER, COBBLER and TD-Close algorithms.

KEYWORDS: Data Mining, Association Rule Mining, High Dimensional Data, Carpenter, Cobbler, TD-CLOSE

I. INTRODUCTION

1.1. Association Rules

Association Rule Mining has become one of the core data mining tasks, and has motivated tremendous interest among data mining researchers and practitioners [1]. It has an elegantly simple problem statement, that is, to find the set of all subsets of items (called itemsets) that frequently occur in many database records or transactions, and to extract the rules telling us how a subset of items influences the presence of another subset. In other words, Association rule mining finds all rules in the database that satisfies some minimum *support* and minimum *confidence* constraints [2]. There are several algorithms have been developed till today. In a database of transaction D with a set of n binary attributes(items) I , a rule defined as an implication of the form $X \rightarrow Y$ where $X, Y \in I$ and $X \cap Y = \emptyset$. where X and Y are called Antecedent and Consequent of the rule respectively. The support, $\text{supp}(X)$, of an item set X is defined as the proportion of transactions in the data set which contain the item set. The confidence of a rule is defined as $\text{Conf}(X \rightarrow Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)}$ [52].

Association rules are defined as an implication rules that informs the user about set of items, which are likely to occur in a transactional data base [1], [2]. They are advantageous to use because they are simple, intuitive and do not make assumptions of any model. Their mining requires satisfaction of user- specified minimum support and confidence from a given data base at the same time. The process of discovering association rules is decomposed into two steps: in step one, generate all the item combinations i.e. frequent itemsets whose support is greater than the user specified minimum support. In step two secondly, use the identified frequent itemsets to generate the rules that satisfy a user

specified confidence. The generation of frequent itemsets requires more effort and the rule generations are discussed.

Association rule mining on high dimensional data, now a days a high topic of research interest in many fields of data mining tasks. There are numerous data repositories which stores the data in different dimensions. Mining association rules on these data bases is a challenging issue. There is a need to find association rules on high dimensional data a more effectively for different applications and decision making. There has been work on quantifying the "usefulness" or "interestingness" of a rule [54]. What is useful or interesting is often application-dependent. The need for a human in the loop and providing tools to allow human guidance of the rule discovery process has been articulated, for example, in [55] [56].

1.2. Outline of the paper

In this paper we deal with algorithmic aspects of association rule mining on high dimensional data. In fact, a broad variety of efficient algorithms to mine association rules have been developed during the last years. We are giving the general review of basic ideas and comparisons behind association rule mining on high dimensional data. Section 2 provides an analysis of high dimensional data. Section 3 represents some detailed analysis of various association rule mining algorithms that are applied on 2-Dimensional, Multi Dimensional and High Dimensional data sets. Section 4 presents brief comparative analyses of assorted algorithms. Section 5 provides conclusions.

II. RELATED WORK

In this paper we mainly restrict ourselves to the classic association rule problem. The mining of all association rules existing in the data base D with respect to certain user interesting measures. D in this case consisting of a data set with varying attributes or dimensions in the average out a total set of 1000-10000 data set. Although the association mining is still topic of further research, during recent years many algorithms for specialised tasks have been developed: First of all, there are the approaches that enhance the association rules itself. E.g. quantitative association rules [19], generalized association rules [31] [32] and to some extent the work on sequential patterns [33][34]. Moreover there are several generalizations of the rule problem [35] [36].

In addition algorithms were developed that mine well defined subsets of the rule set according to specified dimensions or interesting measures etc, e.g. General constraints [37][38], optimized rules[39] [40], maximal frequent itemsets[45], and frequent closed itemsets [46][30]. Moreover there are algorithms to mine dense data bases [42] and speed tables. These approaches are supplemented by algorithms for online mining of association rules [43] and incremental algorithms e.g. [44] [41]

III. BASIC PRELIMINARIES

3.1. Problem formulation

Let T be a discretized data table (or data set), composed of a set of rows, $S = \{r_1, r_2, \dots, r_n\}$, where r_i ($i = 1, \dots, n$) is called a row ID, or rid in short. Each row corresponds to a sample consisting of k discrete values or intervals. For simplicity, we call each of this kind of values or intervals an item. We call a set of rids a rowset, and a rowset having k rids a k -rowset. Likewise, we call a set of items an itemset. A k -rowset is called large if k is no less than a user-specified threshold which is called minimum size threshold. Let TT be the transposed table of T , in which each row corresponds to an item ij and consists of a set of rids which contain ij in T . For clarity, we call each row of TT a tuple. Fig 1 shows an example table T with 4 attributes (columns): A , B , C and D . The corresponding transposed table TT is shown in Table 2.2. For simplicity, we use number i ($i = 1, 2, \dots, n$) instead of r_i to represent each rid.

Originally, we want to find all of the frequent closed itemsets which satisfy the minimum support threshold \min_sup from table T . After transposing T to the transposed table TT , the constraint minimum support threshold for itemsets becomes the minimum size threshold for rowsets. Therefore, the mining task becomes finding all of the large closed rowsets which satisfy minimum size threshold \min_sup from table TT .

rid	A	B	C	D
1	a1	b1	c1	d1
2	a1	b1	c2	d2
3	a1	b1	c1	d2
4	a2	b1	c2	d2
5	a2	b2	c2	d3

Fig 1: an example table T

3.2. Analysis of High Dimensional Data

The emergence of various new application domains, such as bioinformatics and e-commerce, underscores the need for analyzing high dimensional data. Many organizations have enormous amounts of data containing valuable information for running and building a decision making system. Extracting the value of that data is a challenge. First and foremost is to understand and analyze these large amount data for effective decision making. A study on mining large databases is presented in [3]. Generally, in a gene expression microarray data set, there could be tens or hundreds of dimensions, each of which corresponds to an experimental condition. In a customer purchase behaviour data set, there may be up to hundreds of thousands of merchandizes, each of which is mapped to a dimension. Researchers and practitioners are very eager in analyzing these data sets. However, before analyzing the data mining models, the researcher will analyze the challenges of attribute selection, the curse of dimensionality, the specification of similarity in high dimensional space for analyzing high dimensional data set. Association Rule Mining in high dimensional spaces presents tremendous difficulty in generating the rules, much better than in predictive mining. Attribute selection is a one which reduces the impact of high dimensional data space at the time of generating rules.

Dimensionality curse is a loose way of speaking about lack of data separation in high dimensional space [4], [5], and [6]. The complexity of many existing data mining algorithms is exponential with respect to the number of dimensions [4]. With increasing dimensionality, these algorithms soon become computationally intractable and therefore inapplicable in many real applications.

3.3. Dimensionality Reduction

In general, there are two approaches that are used for dimensionality reduction:

- (1) Attribute Transformation
- (2) Attribute Decomposition.

Attribute Transformations are simple function of existent attributes. For sales profiles and OLAP-type data, roll ups as sums or averages over time intervals can be used. In multivariate attribute selection can be carried out by using Principle Component Analysis (PCA) [7].

Attribute Decomposition is a process of dividing data into subsets. Using some similarity measures, so that the high dimensional computation happens over smaller data sets [8]. Dimensions stay the same, but the costs are reduced. This approach targets the situation of high dimensions, large data. It was proven that, for any point in a high dimensional space, the expected gap between the Euclidean distance to the closest neighbour and that to the farthest point shrink as the dimensionality grows [6]. This phenomenon may render many data mining tasks ineffective and fragile because the model becomes vulnerable to the presence of noise.

IV. ALGORITHMS OF ASSOCIATION RULE MINING ON HIGH DIMENSIONAL DATA

In this section, we briefly describe and systematize the most common algorithms. We do this by referring to the fundamentals of frequent itemset generation that was defined in the previous section. Our goal is not to go too much into detail but to show the basic principles and differences between the approaches.

The Association rule mining (ARM) research mainly focuses on discovery of patterns and algorithms. In general, most of the discovered huge numbers of patterns are actually obvious, redundant, and useless or uninteresting to the user. Techniques are needed to identify only the useful/interesting patterns and present them to the user. The data items and transactions are organized in high dimensional space; it is natural to extend mining frequent itemsets and their corresponding association rules to high dimensional space. High dimensional association rules involve more number of dimensions or predicate. In many applications, it is difficult to find strong associations among data items at high dimensions of abstraction due to the sparsity of data and may need of commonsense knowledge representation.

The remainder of this chapter let us see several state-of-art algorithms for mining association rules from high-dimensional data. The discussion starts with various algorithms applied on two dimensional data sets followed by algorithms on multidimensional data sets and finally ends with algorithms applied on high dimensional data.

4.1 Apriori Algorithms

The Apriori algorithm is based on the above-mentioned steps of frequent itemsets and rule generation phases [1], [9]. It is applied on 2-Dimensional data set. Frequent itemsets are generated in two steps. In the first step all possible combination of items, called the candidate itemset (C_k) is generated. In the second step, support of each candidate itemset is counted and those itemsets that have support values greater than the user-specified minimum support from the frequent itemset (F_k). In this algorithm the database is scanned multiple times and the number of scans cannot be determined in advance.

The AprioriTID algorithm is a variation to basic Apriori algorithm [1], [9]. It used to determine the candidate itemsets before each pass begins the main interesting feature of this algorithm is that it does not use the database for support counting after the first pass. In this it uses a set with C_k with the elements Tid and a large k -itemsets present in the transactions with the associated identifier Tid . If a transaction does not contain a k -itemsets then the set will not have any entry for that transaction.

The SETM, the Set Oriented Mining [10], [11]. This uses the SQL join operations for candidate generation [53]. It can also be used k -way join operations for candidate generation [12]. The candidate item set with the TID of generating transaction is stored as a sequential structure, which is used for support counting.

Apriori can be combined with BFS and DFS. In BFS it scans the data base once for every candidate set size k [51]. When using DFS the candidate sets consists only of the item sets of on of the nodes in the tree. The simple combination of DFS with counting occurrences is therefore no practical relevance [50].

4.2. Partition Algorithm

The Partition algorithm differs from the Apriori algorithm in terms of the number of database scans [13]. The partition algorithm scans the database at most twice. The algorithm is inherently parallel in nature and can be parallelized with minimal communication and synchronization between the processing nodes the algorithm is divided into 2 phases: i) during the first phase, the database is divided into n non-overlapping partitions and the frequent itemsets for each partition are generated. ii) In the second phase, all the local large itemsets are merged to form the global candidate itemsets and a second scan of the database is made to generate the final counts.

The partition algorithm is designed to generate rules in parallel and utilize the power of a number of processors. It is used to aggregate the power and memory of many processors. The data in the form of a single file is distributed among different processors with each processor generating itemsets for that part of the data in parallel. This would require the passing of intermediate information among processors to generate the global rules. The parallelized partition algorithm, although developed for multiple processors, can be adapted for multiple database scenarios where data is distributed over multiple databases

4.3. Parallel Mining of Association Rules

Although Apriori is a simplest sequential ARM algorithm designed so far but it has some limitations like large number of candidate itemsets were generated that scans database at every step. To

overcome these demerits several parallel algorithms are used to discover the association rules [14]. In spite of the significance of the association rule mining and in particular the generation of frequent itemsets, few advances have been made on parallelizing association rule mining algorithms [49] [48]. Most of these algorithms are based on Shared memory Multiprocessor (SMP)[14] architecture based on Apriori like algorithms.

These algorithms are designed on different platforms i.e. shared memory system (SMS) and distributed memory system (DMS). CCPD (common candidate partition data base) and PCCD (Partitioned Candidate Common Database) proposed on a shared memory system. Three Apriori based parallel algorithms are used to mine association rules in parallel mining approach. They are as follows:

The *Count Distribution* algorithm is a straight-forward parallelization of *Apriori*. Each processor generates the partial support of all candidate itemsets from its local database partition. At the end of the each iteration the global supports are generated by exchanging the partial supports among all the processors.

The *Data Distribution* algorithm partitions the candidates into disjoint sets, which are assigned to different processors. However to generate the global support each processor must scan the entire database (its local partition, and all the remote partitions) in all iterations. It thus suffers from huge communication overhead.

The *Candidate Distribution* algorithm also partitions the candidates, but it selectively replicates the database, so that each processor proceeds independently. The local database partitions are still scanned at each and every iteration.

4.4. Incremental Mining of Association Rules

The Incremental mining algorithm is used to find new frequent itemsets [15], [16], [17]. It requires minimal recompilation when new transactions are added to or deleted from the transaction database. Incremental mining algorithm is combined with negative border approach to find association rules on high dimensional data attributes using sampling on large databases. The negative border consists of all itemsets that were candidates, which did not have the minimum support [18]. During each pass of the Apriori algorithm, the set of candidate itemsets C_k is computed from the frequent itemsets F_{k-1} in the join and prune steps of the algorithm. The negative border is the set of all those itemsets that were candidates in the k^{th} pass but did not satisfy the user specified support. The algorithm uses a full scan of the whole database only if the negative border of the frequent itemsets expands. Another variant for incremental mining of association rules proposed in [58]. Incremental clustering can be used along with association rule mining to acquire knowledge which can be further under study.

4.5. Multi Dimensional Association Rules

Multi dimensional association rule mining carried on multi dimensional datasets which stores the data in more than one dimensions or predicates [19]. Multidimensional association rules are used to mine when the data or transactions are located in multidimensional space, such as in a relational database or data warehouse. Multiple dimensional association rule mining is to discovery the correlation between different predicates/attributes [20]. Each attribute/predicate is called a dimension. The data may contain different types of data such as categorical, Boolean, numeric. The attributes are also called as quantitative attributes. Association rule mining on these attributes can be carried both in static and dynamic discretization methods [21].

More research work is carried under Dynamic discretization method. In this the numeric attributes are dynamically discretized during the mining process so as to satisfy the mining criteria. Predicate attributes or values in the dimensions are converted into Boolean values so as to carry the mining process very effectively [22]. It is carried on following steps. 1) Determine the number of partitions for each predicate or attribute. 2) Generation of frequent itemsets. 3) Generating interesting rules using minconf. Fabulous research work is carried under Quantitative association rules ranging from low-level predicates to multi dimensional predicate and used for different range of application.

4.6. CARPENTER Algorithm

A new method for finding closed patterns in high-dimensional biological datasets, called CARPENTER [23]. This integrates the advantages of vertical data formats and pattern growth

methods to discover stored patterns in data sets. By converting data into vertical data format {item: TID_set}, the TID_set can be viewed as rowset and the FP-tree so constructed can be viewed as a row enumeration tree. CARPENTER conducts a depth-first traversal of the row enumeration tree, and checks each rowset corresponding to the node visited to see whether it is frequent and closed.

4.7. COBBLER Algorithm

A COBBLER algorithm is used to find frequent closed itemset by integrating row enumeration with column enumeration [24]. Its efficiency has been demonstrated in experiments on a data set with high dimension and a relatively large number of rows. Mining frequent itemsets in the presence of noise in the large data sets is another issue to consider [25]. Since the data set may contain noisy and inconsistent data, which may affect the performance at the time of mining. Several algorithms and analysis work is presented to mine frequent item.

4.8. TD-Close Algorithm

TD-Close used to find the complete set of frequent closed patterns in high dimensional data [26]. It exploits a new search strategy, top-down mining, by starting from the maximal rowset, integrated with a novel row enumeration tree, which makes full use of the pruning power of the min_sup threshold to cut down the search space. A new algorithm for discovering maximal frequent set [59] is combined along with TD-Close to find efficient frequent set.

4.9. Pattern-Fusion Algorithm

Furthermore, an effective closeness-checking method is also developed that avoids scanning the dataset multiple times. Even with various kinds of enhancements, the above frequent, closed and maximal pattern mining algorithms still encounter challenges at mining rather large called *colossal* patterns [27], since the process will need to generate an explosive number of smaller frequent patterns. Colossal patterns are critical to many applications, especially in domains like bioinformatics. A novel mining approach is investigated, called Pattern-Fusion [28], which efficiently finds a good approximation, to colossal patterns. With Pattern-Fusion, a colossal pattern [27], is discovered by fusing its small fragments in one step, whereas the incremental pattern-growth mining strategies, such as those adopted in *Apriori* and *FP-growth*, have to examine a large number of mid-sized ones. Mining partial periodic patterns in time series databases is another interesting study [57]. This property distinguishes Pattern-Fusion from existing frequent pattern mining approaches and draws a new mining methodology. Further extensions on this methodology are currently under investigation. Long patterns series from huge databases can be mined efficiently using [60].

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section we compare the algorithms and explain the observed differences in performance behaviour.

5.1. Experimental Results

To carry out performance studies we implemented the common algorithms to mine association rules starting with Apriori to multi and high dimensional algorithms like CARPENTER and COBBLER algorithms.

To evaluate the efficiency of the algorithms varying with different dimensionality along with the Apriori algorithm is implemented at the same condition. We use different synthetic data sets and real data set sample Basket Sales contain 3 to 100 dimensions, Car Equipment Data, Lung Cancer(LC), ALL, OC, POS, Thrombin database which contains 3 to 15000 dimensions. The commonly used data sets are shown in figure 2 and figure 3. We performed our experiments using a Pentium IV 1, 8 Gigahertz CPU with 512MB.

The experiments in figure 4 to figure 7 were carried out on a synthetic data sets from [2][47]. These data sets were generated with a data generator [2] that simulates the buying behaviour of customers in retail business. Data set "T10.I4.D100K" means an average transaction size of 10, an average size of the maximal potentially frequent itemsets of 4 and 100,000 generated transactions. The number of patterns was set to 2,000 and the number of items to 1,000.

In addition to the experiments from [2][47], we restricted the maximal length of generated itemsets from 1 up to 9 on the data set “T20.I4.D100K” at min_sup= 0.33%, c.f. figure 8, 9, 10 shows the behaviour of the algorithms on real world applications. The basket data consists of about 70,000 customer transactions with approximately 60,000 different items. The average transaction size is 10.5 items. The car equipment data contains information about 700,000 cars with about 6,000 items. In the average of 20 items are assigned to each car. It is ascertaining that all the algorithms scale linearly with the data base size.

Data set	#item	#row	Avg row length
thrombin	13935	1316	29745
POS	1657	595597	164
synthetic data	100000	15000	1700

Fig 2: Data set for COBBLER,

Data set	#item	#row	No. Of. Dimensions
LC	1000	33	12533
ALL	10000	32	2173
OC	1000	253	532

Fig 3: Data set for CARPENTER, and TD_Close

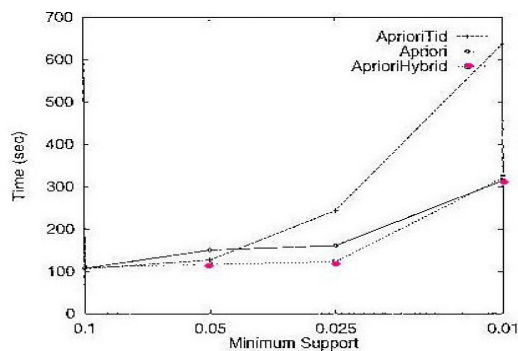


Fig 4: T10.I4.D100K

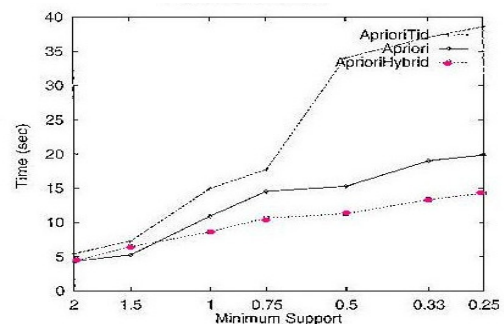


Fig 5: T20.I4.D100K

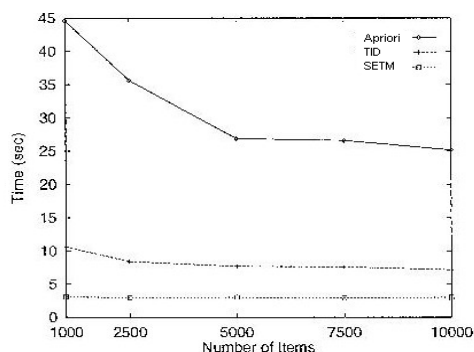


Fig 6: Basket data

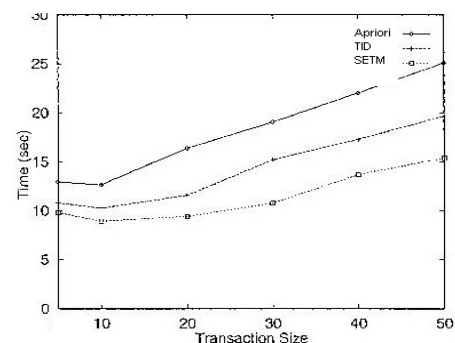


Fig 7: Car Equipment Data

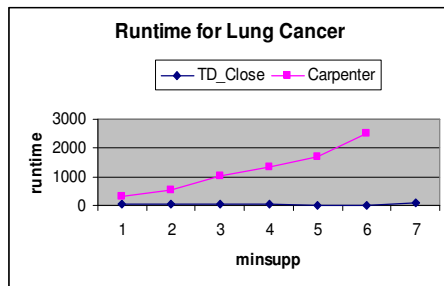


Fig 8: Run time for Lung Cancer

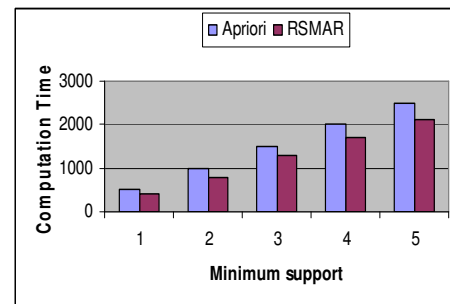


Fig 9: Run time for Sales data with 3 dimensions

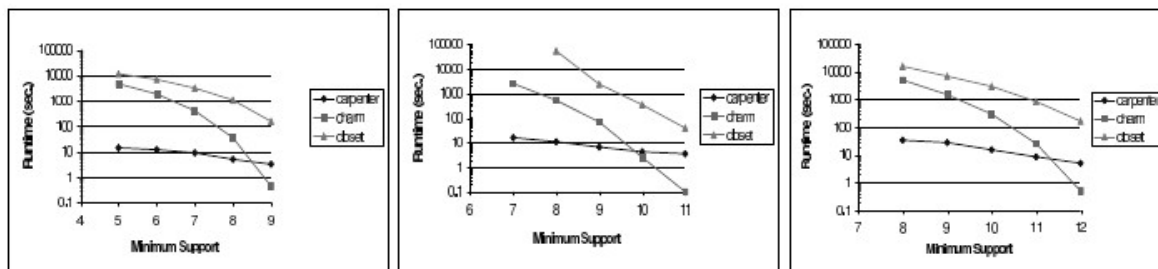


Fig 10: Varying min_sup with l=0.6: a) LC, b) ALL, c) OC

5.2. Comparative Study and Discussion

The several approaches to mine association rules and the need of high dimensional data analysis were discussed on high dimensional data. It represents major issues in analyzing the high dimensional data and their importance. Association rule mining algorithms such as *Apriori* and *AprioriTID* etc on two dimensional data, RSMAR, Fuzzy, and QARM algorithms on Multi dimensional data, CARPENTER, COBBLER, and TD-CLOSE etc algorithms on high dimensional data are used to find patterns form the databases. They strive to provide approximate solutions on low cost and improve the performance. Their performance categorizations are as follows.

The basic *Apriori*, *AprioriTID* [1] [9] and *SETM* [10] algorithms ascertain that they are efficient for mining association rules on two dimensional data. But they suffer with multiple scans over the database and as the database size increases it shows effect on performance.

The Partition Algorithm [13] for mining of Association Rules is better than the basic approach in reducing the number scans and increasing the efficiency when the data base size is increase. Six synthetic datasets of varying complexity are generated. These data sets are varying from very few transactions with few frequent itemsets to larger transactions with more number of frequent itemsets. The Incremental Mining of Association rules are proved efficient scalable over the earlier *Apriori* Algorithms [18]. It is observed that the number of passes or scans over the database is reduced with respect to varying number of rows and frequent patterns. Its performance evaluation is carried out on large databases with various dimensionalities and proved efficient in finding frequent patterns.

The performance of Association rule mining on multidimensional dataset was discussed on RSMAR [22], and QARM [29]. It is observed that RSMAR performs better and more rapid than basic traditional *Apriori*, which decreases the number of database scans and reduces the computations.

CARPENTER is much faster that earlier CHARM [3] and CLOSET [30], in finding frequent closed patterns on data sets with small number of rows and large number of features represented as dimensions. CARPENTER outperforms better in running time and faster by varying the minimum support and minimum length ratio [23]. It is observed that it outperforms better run time at higher support level i.e. as *min_sup* is decreased and much faster in finding frequent patterns.

COBBLER is much faster than CARPENTER in finding frequent closed patterns on a data set with large number of rows and features or dimensions [24]. It performs better in runtime by varying

minimum support and row ratio. COBBLER performs better run time and faster when *min_sup* is relatively high and best when it decreased.

The TD-Close algorithm is much faster than CARPENTER after discretization using entropy based method. This is due to the decrease in *min_sup* as the dimensionality is reduced. This further reduces the runtime of the TD-Close algorithm.

VI. CONCLUSION

The purpose of this article is to present a comprehensive classification of different Association rule mining techniques for high dimensional data. Association rule mining has gained considerable prominence in the data mining community because of its capability of being used as an important tool for knowledge discovery and effective decision making. Association rules are of interest to both database community and data mining users. Since data items and transactions are organized in high dimensional space, it is natural to extend mining frequent itemsets and their corresponding association rules to high dimensional space. Due to the huge accumulation in the data by day to day computations it makes attention to the researcher to study several issues pertaining to rule discovery particularly when data is represented in high dimensional space. As the number of dimensions increase, many Association rule mining techniques begin to suffer from the curse of dimensionality, de-grading the quality of the results.

In this paper, the algorithmic aspects of association rule mining on high dimensional data are presented. From the broad variety of efficient algorithms that were developed are compared the most important one. The algorithms are systemized and analyzed their performance based on both run time experiments and theoretical considerations. . From the above discussion it is observed that the current techniques will suffers with many problems. It is also observed that as the diversity in databases and numerous dimensionalities of the databases, the various algorithms are proven to be efficient in finding frequent patterns in terms of run time as the decrease in *support* values. The COBBLER, CARPENTER and TD-Close are the better and much faster algorithms in finding the frequent patterns over the high dimensional data spaces.

REFERENCES

- [1] Agrawal R., Mannila H., Srikant R., Toivonen H, and Inkeri Verkamo A. "Fast discovery of association rules". *Advances in Knowledge Discovery and Data Mining*, pages 307–328. AAAI Press, Menlo Park, CA, 1995.
- [2] Agrawal, R. and Srikant, R. "Fast algorithms for mining association rules". In: *Proceedings of the 1994 International Conference On Very Large Data Bases (VLDB'94)*, Santiago, Chile, pp 487–499, 1994.
- [3] M J Zaki and C J Hsiao," CHARM- an Efficient algorithm for closed itemset mining, in the proceedings of *SDM 2002*, p 457-473., 2002
- [4] Aggrawal, C.C., Hinnwburg, A., and Keim, D.A. "On the surprising behavior of distance metrics in high dimensional space". *IBM Research report*, RC 21739, 2000.
- [5] Beyer K., Goldstein, J.Ramakrishnan, R., and Shaft, U. "When is nearest neighbor meaningful?" In *Proceedings of the 7th ICDT*, Jerusalem, Israel. 1999.
- [6] Beyer K and Ramakrishnan. "Bottom-up computation of sparse and iceberg cubes". In: *Proceeding of the ACM-SIGMOD 1999 International Conference on Management of Data (SIGMOD'99)*, Philadelphia, PA, pp 359–370, 1999.
- [7] Mardia,K, Kent, J and Bibby,J."Multivariate Analysis". *Academic Press*, 1980.
- [8] McCullum. A., Nigam, K., and Ungar, L.H." Efficient clustering of high dimensional data sets with application to reference matching". In *proceedings of the 6th ACM SIGKDD*, 167-178, Boston., MA, 2000.
- [9] Thomas, S. and S. Chakravarthy. "Incremental Mining of Constrained Associations". In the 7th *International Conference of High Performance Computing (HiPC)*. 1998.
- [10] Maurice Houtsma., Arun Swami., "Set-Oriented Data Mining in Relational Databases", in the Elsevier *proceedings on Data & Knowledge Engineering* , pages 245-262, 1995.
- [11] Han, J., "DMQL: A data mining query language for relational database.". In the *proceedings of ACM SIGMOD workshop on research issues on data mining and knowledge discovery*. Montreal. 1996.
- [12] Mishra, P. and S. Chakravarthy," Evaluation of K-way Join and its variants for Association Rule Mining". *Information and Technology Lab* at the University of Texas at Arlington, TX. 2002.

- [13] Shenoy, P."Turbo-charging Vertical Mining of Large Databases". In *ACM SIGMOD International Conference on Management of Data*. Dallas. 2000.
- [14] Zaki MJ, Parthasarathy S, Ogihara M, Li W, "Parallel algorithm for discovery of association rules". *Data mining and knowledge discovery*, 1:343–374., 1997.
- [15] Thuraisingham, B., " A Primer for Understanding and Applying Data Mining". *IEEE*, 2000. Vol. 2, No.1: p. 28-31.2000.
- [16] Agrawal R." Mining association rules between sets of items in large databases". In: *Proceedings of the ACM-SIGMOD international conference on management of data (SIGMOD'93)*", Washington, DC, pp 207–216, 1993.
- [17] Wei-Guang Teng and Ming-Syan Chen," Incremental Mining on Association Rules".
- [18] H.Toivonen, "Sampling large databases for association rules". In: *Proceeding of the 1996 international conference on Very Large Data Bases (VLDB'96)*, Bombay, India, pp 134–145, 1996.
- [19] Srikant R, Agrawal R "Mining quantitative association rules in large relational table", in the proceedings of *ACM SIGMOD international conference on management of data*, p.1-12 1996.
- [20] Thiwari K., Mitra P., Gupta N., Mangal N .," Mining Quantitative Association Rules in Protein Sequences", in the proceedings of *Data Mining, LNAI 3755*, PP 273-281., 2006.
- [21] Zhao Q and Bhowmick, S S" Association Rule mining – A Survey",
- [22] Anjana P, Kamalraj P,"Rough set model for discovering Multidimensional association rules", in the proceedings of *IJCSNS*, VOL 9, no.6,p 159-164, 2009.
- [23] Pan F, Cong G, Tung AKH, Yang J, Zaki M." CARPENTER: finding closed patterns in long biological datasets". In: *Proceeding of the 2003 30.ACM SIGKDD international conference on knowledge discovery and data mining (KDD'03)*, Washington, DC, pp 637–642, 2003.
- [24] Pan F, Tung AKH, Cong G, Xu X,"COBBLER: combining column and row enumeration for closed pattern discovery". In: *Proceeding of the 2004 international conference on scientific and statistical database management (SSDBM'04)*, Santorin Island, Greece, pp 21–30, 2004.
- [25] Paulsen S, Sun X, Wang W, Nobel A, Prins J." Mining approximate frequent itemsets in the presence of noise: algorithm and analysis". In: *Proceeding of the 2006 SIAM international conference on data mining (SDM'06)*, Bethesda, MD, pp 405–416, 2006.
- [26] Han, Liu, H. D. Xin and Z. Shao. "Mining frequent patterns on very high dimensional data: A top down row enumeration approach.". *Proceedings of the 2006 SIAM International Conference on Data Mining, (ICDM2006)*, Bethesda, MD., pp: 280-291.2006.
- [27] Yan X, Han J, Yu PS, Cheng H . "Mining colossal frequent patterns by core pattern fusion". In: *Proceeding of the 2007 International Conference On Data Engineering (Icde'07)*, Istanbul, Turkey, 2007.
- [28] Zhu .F, Yan.X, Han.J, Philip S Yu. "Mining Colossal Frequent Patterns by core pattern fusion". In *proceeding of 2007 International Conference on Data Engineering*, 2007.
- [29] Aumann, Y. Lindell, Y. "A Statistical Theory for Quantitative Association Rules." *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 261 -270, 1999.
- [30] J Pei., J.Han., R.Mao ,," CLOSET: An Efficient Algorithm for mining Frequent closed Itemsets", in proceedings of the 2000ACM SIGMOD international conference on Management of Data, Dallas, Texas, USA p 21-30., 2000.
- [31] R.Srikant and R.Agrawal ,," Mining generalized association rules", in proceeding of the 21 st conference on very large databases(VLDB'95), Zurich, Switzerland september 1995.
- [32] J Hipp, A Myka, R Wirth and U . Guntzer," a new algorithm for faster mining of generalized association rules", in proceedings fo the 2nd European Symposium on Principles of data mining and knowledge discovery (PKDD'98), Nantes, France, September 1998.
- [33] R Agrawal land R. Srikant ,," Mining Sequential patterns:", in *Proceedings of the internation conference on data Engineering (ICDE)* Taipei, Taiwan, March 1995.
- [34] H Mannila, H Toivonen and I. Verkamo," Discovery of Frequet episodes in event sequences", *Daa mining and Knowledge discovery*, 1(3), November 1997.
- [35] R Motwani, E.Cohen, M. Datar, S.Fujiware, A.Gionix, P.Indyk, J.D. Ullman and C.?Yang, " finding intersitng assoication without support pruning ", in *Proceedings of the 16th international conference on data engineering (ICDE)*, IEEE 2000.
- [36] D.Tsur,J.D.Ullman,S.Abitboul, C.Clifton, R.Motwqani, S.Nestorov, and A.Rosenthal, " Query flocks: A generalization of association rule mining", in proceedings of 1998 ACM SIGMOD international conference on Management of data, Seattle, Washington, USA, June 1998.
- [37] R.Ng, L.S.Lakshmanan, J. Han and A. Pang, "Exploratory mining and pruning optimizations of constrained assoication rules", in proceedings of 1998 ACM SIGMOD international conference on Management of data, Seattle, Washington, USA, June 1998.

- [38] R. Srikant , Q.Vu and R.Agrawal,” Mining association rules with item constraints”, in proceedings of the 3rd international conference on KDD and Data Mining (KDD’97),Newport Beach, California, August 1997.
- [39] T. Fukuda, Y. Morimoto, S.Morishta and T.Tokuyama, “ Mining optimized association rules for numeric attributes”, in proceedings of the 15th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems(PODS’96), Montral, Canada, June 1996.
- [40] R. Rastogi and K.Shim,” Mining optimezed support rules for numeric attribytes”, in proceedings of the 15th international conference on Data engineering, Sydney, Australia, March 1999, IEEE Computer society Press.
- [41] 4N.F. Ayan, A.U.Tansel and E.Arkun,” An efficient algorithm to update large itemets with early pruning”, in proceedings of the 5th international conference on knowledge discovery and data mining(KDD’99), San diego, California, USA, August 1999.
- [42] 5R.J. BAYardo Jr.,R.Agrawal and D.Gunopulos,” Constraint-based rule mining in large, dense databases”, in proceedings of the 15th international conference on Data engineering , Sydney, Australia, March 1999.
- [43] C Hidber,” Online Association Rule mining”, in proceedings fo the 1999 ACM SIGMOD Conference on Management of Data, 19899.
- [44] S.Thomas, S.Bodagala, K Alsabri and S.Ranka ,“ An efficient algorithm for the incremental updaton of association rules in large databases. In proceedings of the 3rd international conference on KDD and Data mining (KDD’97), Newportbeach, California, August 1997.
- [45] M. J.Zaki , S.Parthasarathy ,M.Ogihara, W.Li,” New algorithm for fast discovery of assocition rues”, in proceedings of the 3rd intenational conference on KDD and Data Mining (KDD’97), Newportbeach, California, August 1997.
- [46] N.Pasquier, Y.Bastide,R.Taouil,and L.Lakhal,” Discovering frequent closed itemsets for association rules”, in proceedings of the 7th inernational conference on database theory (ICDT’99), Jerusalem, Israel, January 1999.
- [47] A.Savasere, E.Omiecinski, and S.Navathe. An efficient algorithm for mining associaton rules in large databases”, in the proceedings of the 21st conference on very large databases (VLDB’95), Zurich, Switzerland, September 1995.
- [48] .R. Agrawal and J. C. Shafer. Parallel mining of associationrules: Design, implementation, and experience. *IEEE Trans.Knowledge and Data Engineering*, 8:962–969, 1996.
- [49] D. Cheung, J. Han, V. Ng, A. Fu, and Y. Fu. A fast distributed algorithm for mining association rules. In *Proc.1996 Int. Conf. Parallel and Distributed Information Systems*, pages 31–44, Miami Beach, Florida, Dec. 1996.
- [50] J.Hipp, U.Guntzer, and G.Nakhaeizadeh,” Mining association rules: Deriving a superior algorithm by analyzing today’s approaches”, in the proceedings of the 4th European conference on Priciples and Practices of knoledge Discovery, Lyon, France, September 2000.
- [51] S.Brin, R.Motwani, J.D. Ullman, and S.Tsur. “ Dynamic itemset counting and impliction rules for market basket data”, in the proceedings fo ACM SIGMOD international conference on Management of Data, 1997.
- [52] M. Holsheimer and A. Siebes. Data mining: The search for knowledge in databases. Technical Report CS-R9406, CWI, Netherlands, 1994.
- [53] M. Houtsma and A. Swami. Set-oriented mining of association rules. Research Report RJ 9567, IBM Almaden Research Center, San Jose, Cali-fornia, October 1993. G.
- [54] Piatestsky-Shapiro. Discovery, analy- sis, and presentation of strong rules. In G. Piatestsky-Shapiro, editor, *Knowledge Dis- covery in Databases*. AAAI/MIT Press, 1991.
- [55] R. Brachman et al. Integrated support for data archaeology. In *AAAI-93 Workshop on Knowledge Discovery in Databases*, July 1993.
- [56] R. Krishnamurthy and T. Imielinski. Practitioner problems in need of database research: Research directions in knowledge discovery. *SIGMOD RECORD*, 20(3):76-78, September 1991.
- [57] J.Han, G.Dong, and Y.Yin,” Efficient mining of partial periodic patterns in time series database”, In *ICDE’99* pp.106-115,Sydney, Austrialia, April 1999.
- [58] D. H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2(2), 1987.
- [59] D-I. Lin and Z. M. Kedem. Pincer-search: A new algorithm for discovering the maximum frequent set. In *6thIntl. Conf. Extending Database Technology*, March 1998.
- [60] R. J. Bayardo. Efficiently mining long patterns from databases. In *ACM SIGMOD Conf. Management of Data*,June 1998.

Authors

K. Prasanna is B.Tech and M.Tech. in Computer Science and Engineering. He is a JNIAS Research Scholar and pursuing Ph.D. in same theme from JNTUH-Hyderabad, India. Currently he is working as Assistant Professor in the Department of Computer Science and Engineering, AITS, Rajampet, Kadapa, AP, INDIA. He is having 7 years of teaching experience. His research interests are Data mining, Neural Networks, Network Security and Mobile computing, Data Structures and Algorithms.



M. Seetha, is B.Tech in ECE from Nagarjuna University, Guntur, India in 1992, M.S.(Software Systems) from BITS, Pilani, India in 1999., and Ph.D in Computer Science Engineering discipline from JNTU Hyderabad, India in 2007. She has 16 years of teaching experience. Presently, she is Professor in Computer Science and Engineering Department at GNITS, Hyderabad, India. Her areas of interest are Information Processing, Data Mining, NLP, Neural Networks and Computer Architectures. She has published more than 25 research papers in reputed journals and conferences.



ACHIEVING EFFICIENT LOAD BALANCING IN PEER TO PEER NETWORK

Ritesh Dayama¹, Ranjeet Kagade², Kedar Ghogale³

^{1&2}M.E., Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon, Pune, India

³Centre for Development of Advanced Computing (C-DAC), Pune, India

ABSTRACT

The Internet traffic is growing, and its nature changes because of new applications. Multimedia applications require bandwidth reservations that were not needed initially when the file transfers dominated the Internet. P2P applications are making traffic patterns impossible to predict, and the traffic loads generated at nodes need to be routed regardless of the traffic pattern. Our goal in this paper is to achieve efficient load balancing in P to P network. Today's routing protocol does not provide load balancing in P to P network such as shortest path routing (SPR) protocols used on the Internet today do not maximize the guaranteed node traffic loads, and do not provide scalable and fast bandwidth reservations. Load balancing can improve the network throughput for arbitrary traffic pattern. In this paper we analyze and implement a routing protocol that is based on load balancing and a commonly used shortest path routing protocol, and is, consequently, termed as LB-SPR. LB-SPR is optimized for an arbitrary traffic pattern. Optimization assumes only the weights assigned to the network nodes according to their estimated demands. It will be shown that the optimized routing achieves the throughputs which are significantly higher than those provided by the currently used SPR protocols, such as OSPF or RIP. Importantly, This new methods calculates the guaranteed traffic loads and so allows fast autonomic bandwidth reservations which are the key for the successful support of triple-play applications, including video and audio applications that require high QoS.

KEYWORDS: Shortest path routing (SPR), Open Shortest Path First (OSPF).

I. INTRODUCTION

The Internet traffic has experienced some major changes lately, which require modifications in the network planning and routing protocols. Heavy traffic loads are generated by the multimedia applications, and the actual traffic distribution in the network becomes very hard to predict, due to the developing peer-to-peer services. On the other hand, the traditional approach to traffic grooming and routing optimization in the optical networks assumes that the traffic demands between pairs of nodes are known, which often not the case is. New routing protocols should be able to optimally utilize the network without knowing the actual traffic distribution. It is widely accepted that the next-generation networks should become more autonomic in the process of the network configuration, topology change detection and adaptation to the traffic load changes. Some of these features are incorporated into today's IP networks: they have the ability to detect the topology changes and change the routing accordingly, the TCP congestion control mechanism adapts the transmission speed to the traffic load changes, etc. These applications require high quality of service: bandwidth reservations and delay guarantees. Centralized bandwidth reservations can obviously become a bottleneck in large-scale networks, as well as the reservations which require each router to know about the available link capacities in the whole network. So, a new mechanism for fast bandwidth reservations is needed. Because of the traffic unpredictability, the customers attached to the network nodes should be served

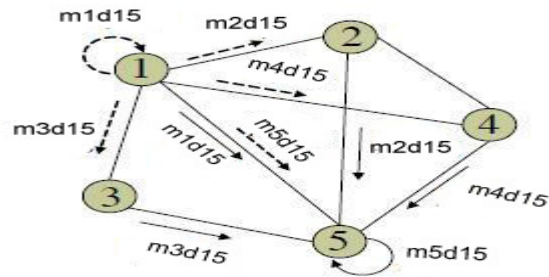
regardless of the traffic pattern between them. In other words, the guaranteed node traffic loads should be sufficient to support all the users attached to these nodes. When the guaranteed node traffic loads are determined, the bandwidth reservations through the network become simple. Each session learns from its router (node) if it can be passed through the network, since the router knows its guaranteed traffic load and the already reserved capacity. If the session can be passed, its request for the bandwidth reservation is passed to the destination router, which checks if there is sufficient capacity on its links toward customers since it knows its guaranteed traffic load and the already reserved capacity. In this way, bandwidth reservations are distributed and are consequently agile. For each session, only two edge routers check their available capacities. And, each router handles bandwidth reservation only for the flows that are either entering or leaving the network through that router. Fast automated bandwidth reservations are very important for growing multimedia applications that demand high QoS, i.e. bandwidth and delay guarantees. If all the flows of equal priority negotiate certain policing interval, the delay guarantees can be achieved when the bandwidth is reserved.

II. LOAD BALANCED SHORTEST PATH ROUTING (LB-SPR)

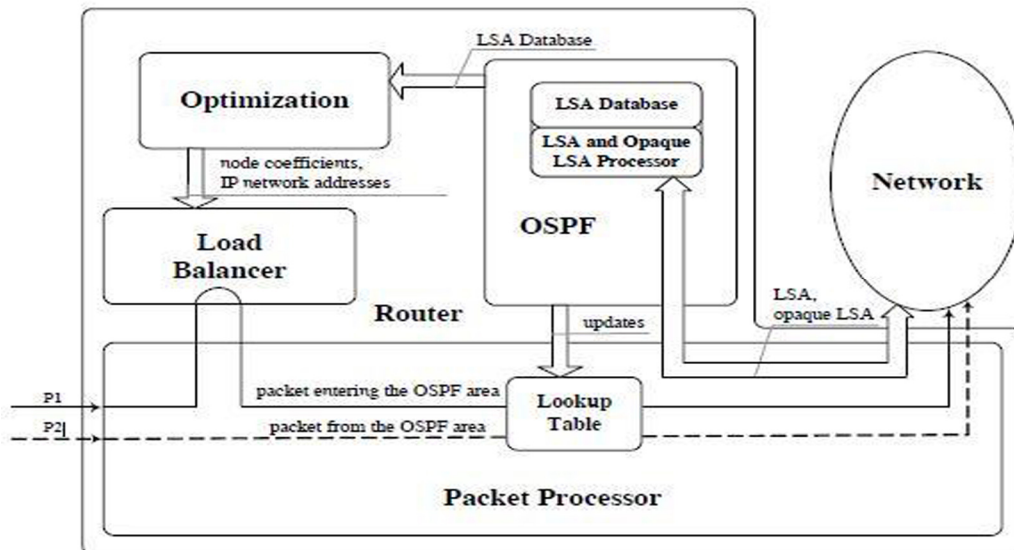
As already described, the proposed routing strategy uses the standard OSPF combined with load balancing, to route the traffic between a pair of nodes in two phases. It distributes the load more evenly among all the links in the network, thereby lowering the average link utilization for congested links, and avoiding bottlenecks. The routing algorithm was proposed, that uses load balancing, and the traffic between every pair of nodes in the network is routed in two phases. First, portions of the routed flow are directed to the balancing routers, according to the balancing coefficients assigned to the routers in the network. Then, in the second phase, every balancing router sends the traffic to its final destination. In this each phase uses the standard shortest path routing (SPR) protocol. In LB-SPR, every packet is routed in two phases, with SPR as the underlying routing protocol in both of the phases.

When a packet arrives to the source router, its intermediate router is determined. The packet is sent to the intermediate router using the standard SPR protocol, and from the intermediate router to the destination router again using the standard SPR protocol. The load is balanced across the intermediate routers, meaning that the specified portions of each flow are transmitted through the intermediate routers. These portions are referred to as the balancing coefficients. A balancing coefficient depends only on the associated balancing router. Balancing coefficients are optimized to maximize the network throughput while ensuring that nodes can generate and receive loads which are proportional to the allocated weights. The node weights are chosen to reflect the expected demands at the nodes. The LB-SPR protocol uses the signaling of the OSPF protocol. Through this signaling, each router in the network is learning the network topology, and the capacity of the nodes' external (customer) links. The external link capacities are taken to be the node weights. The OSPF signaling had to be extended, to distribute the information about the link capacities, as well. Based on the information provided through the OSPF signaling, the OSPF routing tables are calculated and the routing optimization is performed. The optimal values of the balancing coefficients are determined for all the routers, using linear programming. Now, the packets are routed based on the balancing coefficients, using the standard OSPF and the loose source routing. Consequently, LB-SPR maintains autonomic fault recovery mechanism developed within OSPF. Namely, whenever there is a network topology change, the routing is adjusted accordingly.

In the proposed routing scheme, the traffic between a node pair (i, j) is routed in two phases. First, portions of the flow from i to j are routed to the intermediate nodes $m \in V$ (V is the set of network nodes). In the next phase, every intermediate node forwards the traffic to its final destination j . The traffic from i to m , and from m to j is routed along the shortest paths. The portion of the flow that is balanced across node m equals k_m , and does not depend on i and j . Of course, $\sum_{m \in V} k_m = 1$. Fig. 1 illustrates the case of routing the traffic between the nodes 1 and 5. The first phase of the flow routing is represented by the dashed arrows, and the second phase of the flow routing by the solid ones.



III. SHORTEST PATH ROUTING



Vol. 2, Issue 1, pp. 455-462

marked the destination as visited (as is the case with any visited intersection) you have determined the shortest path to it, from the starting point, and can trace your way back, following the arrows in reverse.

IV. IMPLEMENTATION

In this section, the implementation of the previously analyzed LB-SPR routing protocol. In order to make LB-SPR as compatible as possible to OSPF, it is implemented in each OSPF area separately. When a packet enters the OSPF area, its intermediate router is determined. The proposed routing scheme uses OSPF to route the packets between the source router and the intermediate router, as well as between the intermediate router and the destination router. Here, the source router is the first router that the packet encounters when it enters the OSPF area, and the destination router is the last router that the packet passes in the OSPF area under consideration. In a common IP router that uses the OSPF protocol, when a packet arrives to the router, it is first processed by the packet processor. The packet processor uses its lookup table to determine the router output port to which the packet should be forwarded based on its destination IP address. The lookup table is updated whenever the network topology changes, which provides an autonomic reliability. A software module calculates new lookup table based on the LSA (Link State Advertisement) control packets exchanged through the OSPF protocol, and sends it to the packet processor. The balancing coefficients are recalculated whenever the network topology changes, which provides the same autonomic reliability as does the OSPF.

The LB-SPR implementation is illustrated in Fig. 2. The solution is based on the OSPF implementation, which is extended to support load balancing. First, it was necessary to allow the retrieval and distribution of the specific information needed by the linear program for the routing optimization, such as the node weights C_i . Finally, the load balancer was implemented to route the packets entering the OSPF area according to LB-SPR. Load balancer first has to determine the intermediate router for each incoming packet, and then to direct the packet accordingly. Specified portions of all the flows entering source routers have to be directed to the intermediate routers, according to the calculated optimal values of the coefficients k_i . We chose the loose source routing as the simplest IP-based solution. Namely, the destination IP address of a packet entering the OSPF area is replaced with the IP address of the intermediate router, while the destination address becomes part of the loose source routing option field.

Let us summarize how the packets are processed in the router shown in Fig. 2. The path for the “new” packet entering the OSPF area is represented with the full line in Fig. 2. The packet which is entering the OSPF area has to be processed by the load balancer, which determines the intermediate router for the packet, and modifies the IP header accordingly. Once the packet has been modified by the load balancer, it is forwarded through the network using the standard OSPF routing tables. On the other hand, the path of the “old” packet that has already been modified by its source router is represented by the dashed line. This packet is only passing through the given router, and does not need to be processed by the load balancer. The information needed to route this packet can be obtained from the standard OSPF routing table.

4.1 Extended OSPF Module

In the case of the regular OSPF, the changes of the network topology trigger the recalculation of the OSPF routes. For LBSPR, every time the topology changes it is also necessary to repeat the routing optimization and recalculate the balancing coefficients k_i , based on the updated OSPF routing tables. The node weights C_i are needed to run the optimization. These weights can be set by the administrator, or can be, more desirably, autonomic. Therefore, we use the SNMP protocol to detect the operational state of the router interfaces in the network, as well as their speeds. The use of the SNMP to detect the changes of the interface operational states (up or down), and their capacities allow together with the OSPF mechanism full automation of the topology change discovery and distribution. Using SNMP, each router learns the operational state of its interfaces and their speeds, and distributes this control information inside the OSPF area. The opaque LSAs with the area-local scope are used to convey this information according to the OSPF standard. Opaque LSAs were introduced to provide a generalized mechanism to allow for the future extensibility of OSPF. Opaque LSA consists of the standard LSA header followed by the 32-bit application-specific information field. In our

implementation, the opaque type value is selected from the range reserved for experimental and private use. The routers' weights, i.e. external link capacities, are transferred as the 64-bit integer values. Incoming and outgoing opaque LSAs are processed and stored into the LSA database.

Whenever the external link capacity changes, the router learns about the change through the SNMP protocol, and distributes the updates by the opaque LSAs. Using this information, the OSPF module calculates the IP routing table and sends this table to the packet processor of the router. Whenever the network topology changes, the OSPF module recalculates the IP routing table and sends its updates to the packet processor. In the LB-SPR implementation, the selected information about the network topology and the capacity of the routers' external (customer) links is transmitted to the optimization module. The OSPF obtains this information from standard LSAs and opaque LSAs. Using this information, the optimization module determines the parameters required to perform load balancing.

4.2 Optimization module

The optimization module gets the required information from the OSPF module which performs the signaling, as we have described in the previous subsection. Based on this information, it optimizes the routing based on load balancing, and sends the required parameters to the load balancer which performs the actual routing of incoming packets.

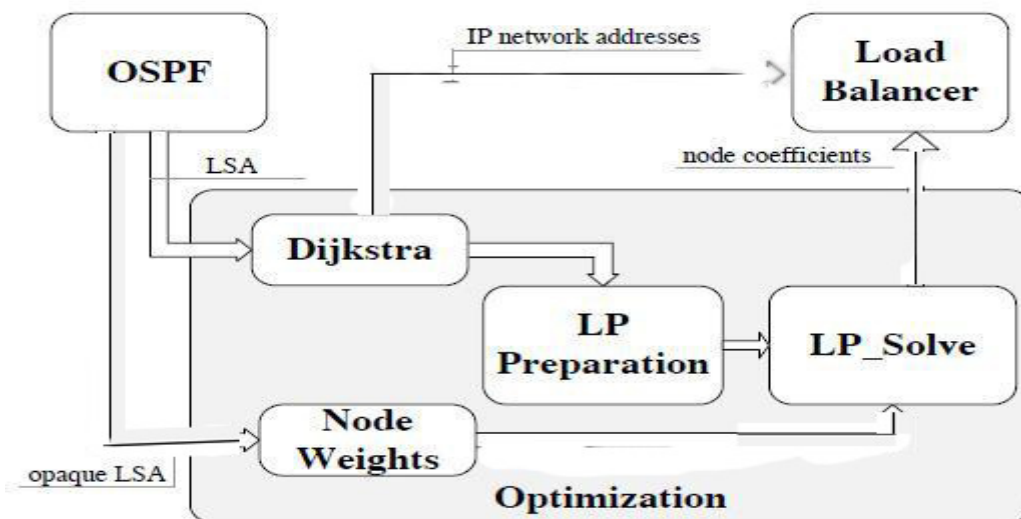


Fig. 3: The scheme of the optimization module [1]

The optimization module is shown in Fig. 3. Based on the network topology information obtained from the OSPF module, the Dijkstra module calculates forwarding trees for all nodes in the network according to the Dijkstra algorithm. The Dijkstra module also calculates the IP network address of each intermediate router through which the traffic will be balanced. This IP address will replace the destination IP address when the source routing is used in the load balancer. Using the calculated trees, the next module in line, the LP preparation module calculates coefficients F_{ij}^l , $i, j \dots V$ which are required for the linear program. Finally, the LP Solve module optimizes the routing and calculates the balancing coefficients k_i , $i \dots V$, which are necessary to the load balancer.

4.3 load balancer

The load balancer receives the balancing coefficients from the optimization module. It also receives the information about the IP network addresses of the intermediate routers. These addresses are calculated by the Dijkstra module, which is the part of the optimization module. The load balancer gets the information that it requires through a TCP connection. Based on this information, the load balancer determines the router output port for each packet entering the router and the OSPF area, and modifies its header in order to balance the traffic appropriately.

For each destination router j , the load balancer of the given source router i stores the information about the currently used intermediate router m_j . We will call router m_j the active intermediate router. It

also maintains a counter with the number of bytes B_j that remain to be balanced across that intermediate router. The initial value of the counter is proportional to the balancing coefficient k_{mj} of the intermediate router m_j . When a packet enters the OSPF area, it has to be processed by the load balancer. First, the destination router for the packet is determined, based on the IP address of the packet destination. Let us say that it is a destination router j . Then, the corresponding IP network address of m_j is found, as well as the counter B_j by the search through a Patricia tree. The Patricia tree allows for a fast lookup. Then, the packet header is modified: the destination address is replaced by the IP network address of the intermediate router m_j , and the original destination address is placed in the option field for the loose source routing. The counter B_j is then decremented by the length of the packet (in bytes). When the counter B_j is smaller than the packet length, the active intermediate router is updated. The next router from the list of possible intermediate routers, $m_j = \text{next}(m_j)$, becomes active, and the counter B_j is set to the value proportional to the balancing coefficient corresponding to that intermediate router, k_{mj} .

V. TESTING

5.1 Testing environment

The network simulation using virtual machines as the network nodes is gaining more popularity with the increase of the computer processing power. The advantage of this simulation method is that the protocol under consideration is executed as it would be executed in the exploitation conditions. This way, the more accurate simulations are performed, and possible errors when adapting the protocol implementation to the simulation tools are avoided. We selected Xen as the virtualization platform because of its maturity and performance [9]. Xen allows several operating systems to be executed on the same computer hardware concurrently. Administrator starts these operating systems through the main operating system, named *dom0*, which boots automatically.

The virtual machines implement the entire TCP/IP stack in which the OSPF protocol will be replaced by the LB-SPR protocol shown in Fig. 2. The virtual machines will play the role of both routers and hosts in the network. The data will be exchanged in the network, and we will test the LB-SPR functioning using the Multi Router Traffic Grapher (MRTG) packet analyzer [10].

5.2 Functional verification of the L LB-SPR implementation

The performance of LB-SPR was analyzed in the network represented in Fig. 4. This is, in fact, the simplified version of the Exodus network topology. For the purpose of this simulation, all the nodes in one city were represented by a single node, and the equivalent link weights were calculated. This network was emulated using seven computers and one Ethernet switch as represented in Fig. 5. Depending on the processor speed and RAM size, the number of the virtual routers executed on a single computer ranges from two to five. The virtual routers X and Y on a single computer are connected through the Xen bridge Xenbr XY.

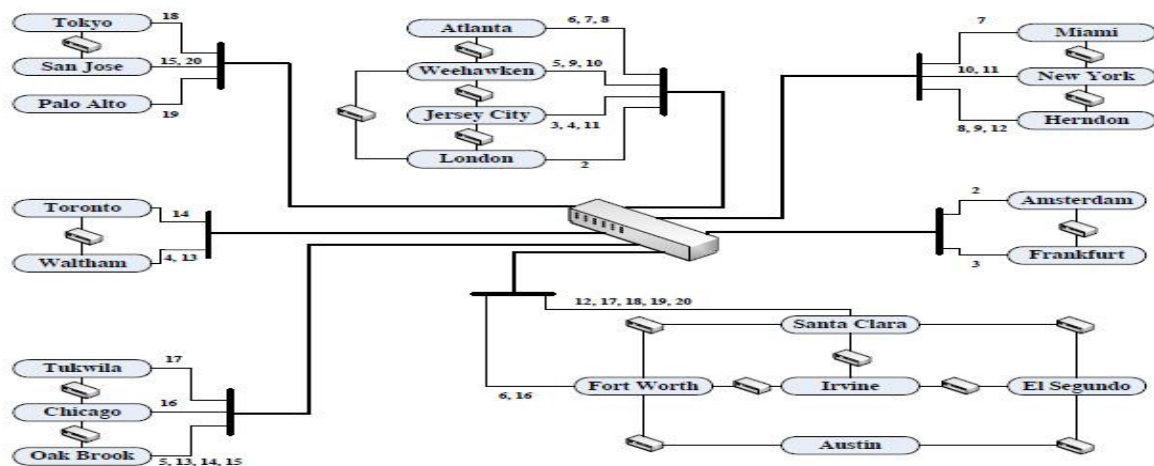


Fig. 4: The simulation environment [1]

Each virtual router is configured using the configuration script. For the analyzed network, the worst-case traffic pattern for OSPF was determined using the maximum matching algorithm. The critical link for OSPF is the link between Tukwila and Santa Clara. It gets congested when the following pairs of nodes communicate with the maximum speeds: Oak Brook - San Jose, Toronto - Palo Alto, Amsterdam - Santa Clara, Tukwila - Irvine, Chicago - Tokyo, and Waltham - El Segundo. The traffic between these nodes was set to the value that causes the critical link utilization to be 100%. Then, the LB-SPR is applied for the same traffic pattern.

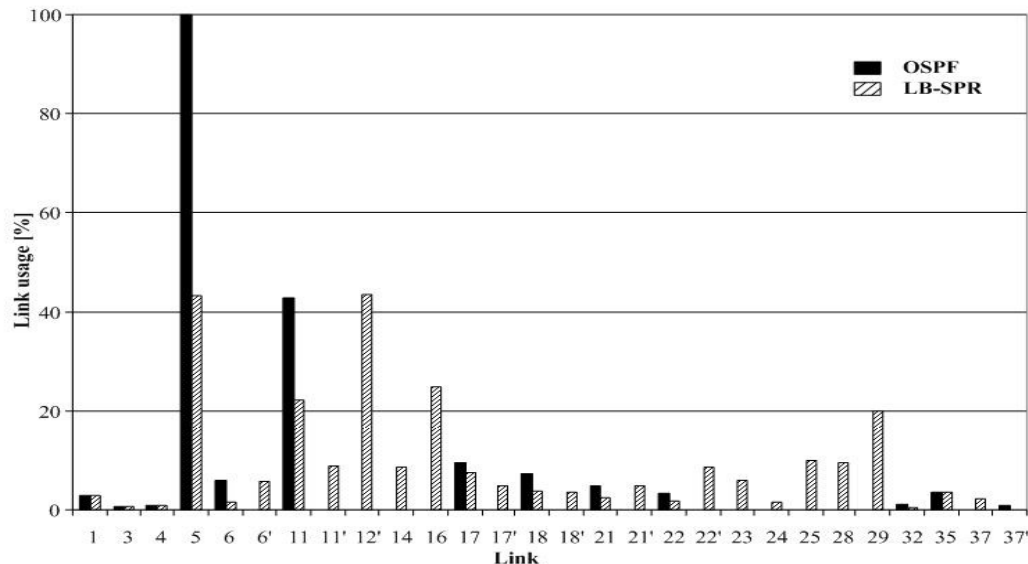


Fig 5: Link Usage [1]

The link utilizations for both OSPF and LB-SPR are observed using the MRGT packet analyzer and plotted in Fig. 5. The link utilizations in the case of OSPF are represented by the black lines and in the case of LB-SPR by the gray ones. It can be observed that the OSPF protocol uses 13 links to route the given traffic pattern, while the LB-SPR protocol uses 27 links. At the same time, the congestion of the critical link is lowered to 43.3% when LB-SPR is applied. Simulation results agreed with the analytical results which confirmed the correctness of the implementation. They show that LB-SPR balances the traffic over the network links better than OSPF. If the critical link 5 had two times lower capacity, LB-SPR would still pass the traffic, while SPR would not.

VI. CONCLUSION

This protocol is automated as the existing routing protocols such as OSPF, and adapts to the changes of the network topology. LBSPR calculates the traffic loads that the nodes can guarantee to carry. Using the information about the guaranteed node traffic loads, the bandwidth reservations become simple in such a network, and, consequently can be made fast. Fast and autonomic bandwidth reservations are important for the multimedia applications whose popularity is growing. At the same time, the LB-SPR protocol maximizes the node traffic loads that can be guaranteed in the given network. It was shown that LB-SPR improves the guaranteed traffic up to 7.7 times for the real networks that we considered, compared to the shortest path routing protocols such as OSPF.

Since LB-SPR is using the OSPF signaling, it inherits its recovery speed which is insufficiently low for the interactive applications. If a faster recovery mechanism is needed, it can be employed at the lower layers as it is typically done. Alternatively, the capacities can be over provisioned to account for the failures to compare the costs of the networks using LB-SPR and OSPF in which the link capacities are over provisioned to pass given node traffic loads even when single failures, of nodes or links, occur.

REFERENCES

- [1] Marija Antić, Nataša Maksić, Petar Knežević, and Aleksandra Smiljanić, "Two Phase Load Balanced Routing using OSPF" IEEE Journal on selected areas in Communications, vol. 28, No. 1,

- January 2010.
- [2] Maksic, N.; Knezevic, P.; Antic, M.; Smiljanic, A.; "On the performance of the load balanced shortest path routing" IEEE Pacific Rim Conference, 2009.
 - [3] M. Antić, A. Smiljanić, "Oblivious Routing Scheme Using Load Balancing Over Shortest Paths", in Proc. ICC 2008, 2008.
 - [4] M. Antić, A. Smiljanić, "Routing with Load Balancing: Increasing the Guaranteed Node Traffics", IEEE Commun. Lett., June 2009.
 - [5] N. Wang, K. Ho, G. Pavlou, and M. Howarth, "An Overview of Routing Optimization for Internet Traffic Engineering", IEEE Commun. Surveys & Tutorials, 1st Quarter 2008.
 - [6] Y. Azar, E. Cohen, A. Fiat, H. Kaplan, and H. Racke, "Optimal Oblivious Routing in Polynomial Time," Proc. of the 35th ACM Symposium on the Theory of Computing, 2003.
 - [7] R. Cohen and G. Nakibli, "On the Computational Complexity and Effectiveness of N-hub Shortest-Path Routing", IEEE Trans. Netw., June 2008.
 - [8] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson, "Inferring Link Weights Using End-to-End Measurements," Proc. of the IMW 2002, Nov. 2002
 - [9] Xen Users' Manual, Xen V2.0 for x86, [Online]. Available: [http://www.xen.org/files/xen user manual.pdf](http://www.xen.org/files/xen%20user%20manual.pdf)
 - [10] MRTG 2.16.2 configuration reference, [Online]. Available:<http://oss.oetiker.ch/mrtg/>

Authors

Ritesh Dayama was born in Solapur, India, in 1988. He received the Bachelor in Computer Science & Engineering degree from the Solapur university of Solapur, in 2009 and the Master in Computer Network degree from the Pune university of Pune, (pursuing), both in Computer engineering. His research interests include Computer Networking, Artificial Intelligence.



Ranjeet Kagade was born in Solapur, India, in 1988. He received the Bachelor in Computer Science & Engineering degree from the Solapur university of Solapur, in 2010 and the Master in Computer Network degree from the Pune university of Pune, (pursuing), both in Computer engineering. His research interests include Computer Networking, Image Processing.



Kedar Ghogale was born in Paranda, India, in 1975. He received the Bachelor in Computer Science & Engineering degree from the Dr.B.A.M. university of Aurangabad, in 1999 and the Master in Computer Network degree from the Pune university of Pune, (pursuing), both in Computer engineering. His research interests include Grid Computing, Genetic Algorithms, Wired Technology, Cloud Computing.

ANALYSIS AND SIMULATION OF SERIES FACTS DEVICES TO MINIMIZE TRANSMISSION LOSS AND GENERATION COST

M. Balasubba Reddy¹, Y. P. Obulesh² and S. Sivanaga Raju³

¹Deptt. of Electrical Engg., Prakasam Engineering College, Kandukur, A.P., India.

²Deptt. of Electrical Engg., L.B.R. College of Engineering, Mylavaram, A.P., India.

³Deptt. of Electrical Engg., J.N.T.U. College of Engineering, Kakinada, A.P., India.

ABSTRACT

This paper presents the severity of over load index (SOL) technique to find the optimal location of FACTS controller to achieve optimal power flow solution. The proposed method estimates the SOL index of transmission line to locate the FACTS device. Main objective of optimal power flow is to minimize the overall cost function which includes total active and reactive power production cost under constraints. Among various FACTS devices thyristor controlled phase shifters (TCPST), thyristor controlled series compensator (TCSC) and static synchronous series compensators (SSSC) are considered to control active power flow through transmission line and also to reduce active power loss. Different operating conditions of the power system are considered for finding the optimal location of FACTS controllers. The proposed technique is an effective method for the optimal location of FACTS controllers.

KEYWORDS: FACTS, TCPST, TCSC, SSSC, OPF

I. INTRODUCTION

In today's highly complex and interconnected power systems, there is a great need to improve electric power utilization while still maintaining reliability and security. Demand of electric power is continuously rising at a very high rate due to rapid industrial development. To meet this demand it is essential to raise the transmitted power along with the existed transmission facilities. The need for the power flow control in electrical power system is thus evident. Power flow is a function of transmission line impedance, the magnitude of the sending end and receiving end voltages and the phase angle between voltages. By controlling one or a combination of the power flow arrangements, it is possible to control the active, as well as, the reactive power flow in the transmission line [1]. The FACTS devices are capable of changing the system parameters in a fast and effective way. It is known that the benefits brought by FACTS devices include improvement of system stability, enhancement of system reliability, and reduction of operation and transmission investment cost [2]. In previous work, researches concentrated on locating and sizing of different types of FACTS devices in order to maximize the power transfer considering networks with variable loads. The problem was formulated as an optimization problem and was solved using different methods such as using iterative techniques, MATLAB optimization routines or Genetic Algorithm (GA) [3-5]. Optimization problem with the objective of minimizing the generating cost in a network with unchanged loads. The problem is solved using OPF algorithm using NR method. Which is fast and simple compared with conventional technique and also give promising results. Improvements of results with FACTS devices is compared with convention N-R OPF method without FACTS devices. OPF is a very large, non-linear mathematical programming problem, the main purpose of OPF is to determine the optimal operation state of a power system while meeting some specified constraints. Since the OPF solution

was introduced by squire [6]. The focus in this paper lies on thyristor controlled phase shifting transformers (TCPST), thyristor controlled series compensators (TCSC) and static synchronous series compensators (SSSC). These facts devices TCPST, TCSC and SSSC by controlling the power flows in the network can help to reduce the flows in heavily loaded lines, to reduce the system loss, reduced cost of production and fulfilled contractual requirements [7]. Location of FACTS devices in the power system are obtained on the basis of static and/or dynamic performance. There are several methods for finding the optimal location of FACTS devices but this paper introduces severity of overload index technique gives optimal location by estimating overload index of each transmission line. The proposed algorithm is tested using the IEEE 5 bus and IEEE 30 bus system and the results are presented.

The description of remaining sections of this paper is as follows. Section 2 introduces OPF without FACTS devices. Modelling of FACTS devices is described in section 3. In section 4 problem formulations is presented. The experimental results on the IEEE5 bus and IEEE30 bus systems are presented in section 5. Finally section 6 summarizes the main conclusion and future scope of the paper.

II. OPF WITHOUT FACTS DEVICES

The objective of active power optimization is to minimize production cost while observing the transmission line and generation active and reactive power limits.

Minimize

$$F_T = \sum_{i=1}^m C_i(P_{Gi}) \quad (1)$$

Subjected to

$$\sum_{i=1}^m P_{Gi} - \sum_{k=1}^n P_{Dk} - P_L = 0 \quad (2)$$

$$P_L \leq P_L^{\max} \quad (3)$$

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max} \quad (4)$$

Where n is the number of system buses and number of generating units respectively. $C_i(P_{Gi})$ is production cost of the unit at i^{th} bus, F_T is the total production cost of m generators, P_{Gi}^{\min} & P_{Gi}^{\max} are minimum and maximum active power limits of the unit at i^{th} bus. P_{Dk} is the active power load at bus k, P_L is the network active power loss, P_l , P_l^{\max} are the active power flow and its limit on line l.

The augmented lagrangian is, [18]

$$L(P_{Gi}) = F_T(P_{Gi}) + \lambda \left(\sum_{k=1}^n P_{Dk} + P_L - \sum_{i=1}^m P_{Gi} \right) + \quad (5)$$

$$\sum_{l=1}^{N_l} \mu_l (P_l - P_l^{\max}) + \sum_{i=1}^m \left[\mu_i^{\max} (P_{Gi}^{\max} - P_{Gi}) + \mu_i^{\max} (P_{Gi} - P_{Gi}^{\max}) \right]$$

λ is for power balance equation.

μ_i^{\min} and μ_i^{\max} are lower and upper active power limits of unit at i^{th} bus.

μ_l is for active power flow limit on line l.

N_l is the number of transmission line flow violations

III. MODELLING OF FACTS DEVICES

In this paper three facts devices are used for comparing the results these are thyristor controlled phase shifting transformers (TCPST), thyristor controlled series compensators (TCSC) and static synchronous series compensators (SSSC). Modelling of these devices is given in the following sections.[20]

3.1 Modelling of TCPST

TCPST can be modelled by a phase shifting transformer with control parameter ϕ . The power flow equations of the line can be derived as follows [15] & [17]

$$P_k = V_k^2 G - V_k V_m [G \cos(\theta_k - \theta_m - \phi) + B \sin(\theta_k - \theta_m - \phi)] \quad (6)$$

$$Q_k = -V_k^2 B - V_k V_m [G \sin(\theta_k - \theta_m - \phi) - B \cos(\theta_k - \theta_m - \phi)] \quad (7)$$

$$P_m = V_m^2 G - V_m V_k [G \cos(\theta_m - \theta_k - \phi) + B \sin(\theta_m - \theta_k - \phi)] \quad (8)$$

$$Q_m = -V_m^2 B - V_m V_k [G \sin(\theta_m - \theta_k - \phi) + B \cos(\theta_m - \theta_k - \phi)] \quad (9)$$

The injected real and reactive power of TCPST at bus k and bus m are as follows.

$$P_{ks} = -V_k^2 G_{ij} \tan^2 \phi - V_k V_l \tan \phi [G_{kl} \sin \delta_{kl} - B_{kl} \cos \delta_{kl}] \quad (10)$$

$$P_{ls} = -V_k V_l \tan \phi [G_{kl} \sin \delta_{kl} + B_{kl} \cos \delta_{kl}] \quad (11)$$

3.2 Modelling of TCSC

It was derived by examining the voltages and currents in the TCSC circuit under the full range of operating conditions. The basic equation is,

$$Z_{TCSC} = R_{TCSC} + jX_{TCSC} = \frac{V_{TCSC}}{I_{TCSC}} \quad (12)$$

Where, V_{TCSC} is the fundamental frequency voltage across the TCSC module,

I_{TCSC} is the fundamental frequency line current and Z_{TCSC} is the TCSC impedance.

Effective TCSC reactance X_{TCSC} with respect to α is,

$$X_{TCSC}(\alpha) = -X_c + C_1(2(\pi - \alpha) + \sin(2(\pi - \alpha))) -$$

$$C_2 \cos^2(\pi - \alpha)(\omega(\tan(\omega(\pi - \alpha))) - \tan(\pi - \alpha)) \quad (13)$$

3.3 Modelling of SSSC

The SSSC OPF model presented in this section enables very flexible and reliable power system optimization studies to be carried out. The flexibility stems from the generality of the SSSC model and the robustness from the strong convergence exhibited by the OPF solution using Newton's method. The SSSC model may be set to control active powers as well as nodal voltage magnitude, at either the sending or the receiving end bus. The mathematical modelling will be obtained from [8] and that was taken as a reference and implemented in the OPF by using the Lagrangian multiplier concept. A SSSC [9] [10] usually consists of a coupling transformer, an inverter and a capacitor. As shown in Fig. 1 the SSSC is series connected with a transmission line through the coupling transformer.

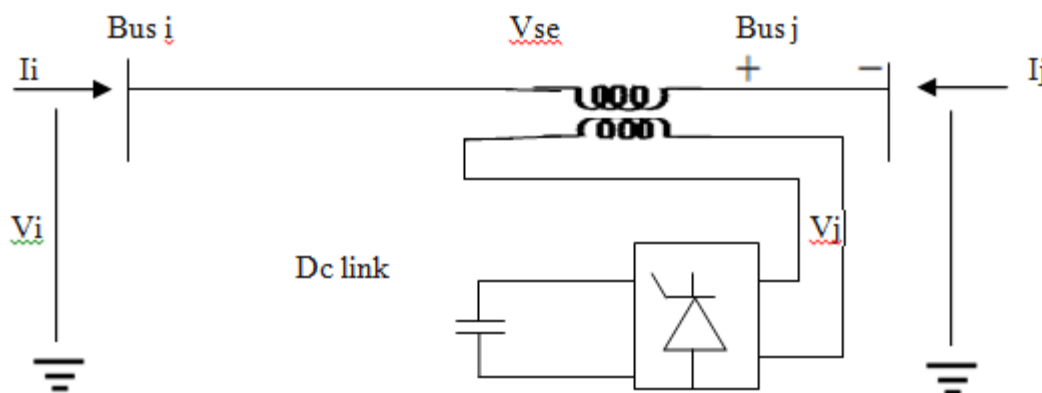


Fig. 1 SSSC basic operating model

It is assumed here that the transmission line is series connected via the SSSC bus j . The active and reactive power flows of the SSSC branch i - j entering the bus j are equal to the sending end active and

reactive power flows of the transmission line, respectively. In principle, the SSSC can generate and insert a series voltage, which can be regulated to change the impedance (more precisely reactance) of the transmission line. In this way, the power flow of the transmission line or the voltage of the bus, which the SSSC is connected with, can be controlled.

An equivalent circuit of the SSSC as shown in Fig. 1 can be derived based on the operation principle of the SSSC. In the equivalent, the SSSC is represented by a voltage source V_{se} in series with transformer impedance. In the practical operation of the SSSC, V_{se} can be regulated to control the power flow of line $i-j$ or the voltage at bus i or j [11].

In the equivalent circuit, $V_{se} = V_{se} \angle \theta_{se}$, $V_i = V_i \angle \theta_i$, $V_j = V_j \angle \theta_j$ then the bus power flow constraints of the SSSC are,

$$P_{ij} = V_i^2 g_{ii} - V_i V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) - V_i V_{se} (g_{ij} \cos(\theta_i - \theta_{se}) + b_{ij} \sin(\theta_i - \theta_{se})) \quad (14)$$

$$Q_{ij} = -V_i^2 b_{ii} - V_i V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) - V_i V_{se} (g_{ij} \sin(\theta_i - \theta_{se}) - b_{ij} \cos(\theta_i - \theta_{se})) \quad (15)$$

$$P_{ji} = V_j^2 g_{jj} - V_i V_j (g_{ij} \cos \theta_{ji} + b_{ij} \sin \theta_{ji}) + V_j V_{se} (g_{ij} \cos(\theta_j - \theta_{se}) + b_{ij} \sin(\theta_j - \theta_{se})) \quad (16)$$

$$Q_{ji} = -V_j^2 b_{jj} - V_i V_j (g_{ij} \sin \theta_{ji} - b_{ij} \cos \theta_{ji}) + V_j V_{se} (g_{ij} \sin(\theta_j - \theta_{se}) - b_{ij} \cos(\theta_j - \theta_{se})) \quad (17)$$

where,

$g_{ij} + jb_{ij} = \frac{1}{z_{se}}$, $g_{ii} = g_{ij}$, $b_{ii} = b_{ij}$, $g_{jj} = g_{ij}$, $b_{jj} = b_{ij}$ The operating constraint of the SSSC (active power exchange via the DC link) is,

$$PE = \text{Re}(V_{se} I_{ji}^*) = 0 \quad (18)$$

where,

$$\text{Re}(V_{se} I_{ji}^*) = V_i V_{se} (g_{ij} \cos(\theta_i - \theta_{se}) - b_{ij} \sin(\theta_i - \theta_{se})) - V_j V_{se} (g_{ij} \sin(\theta_j - \theta_{se}) - b_{ij} \cos(\theta_j - \theta_{se})) \quad (19)$$

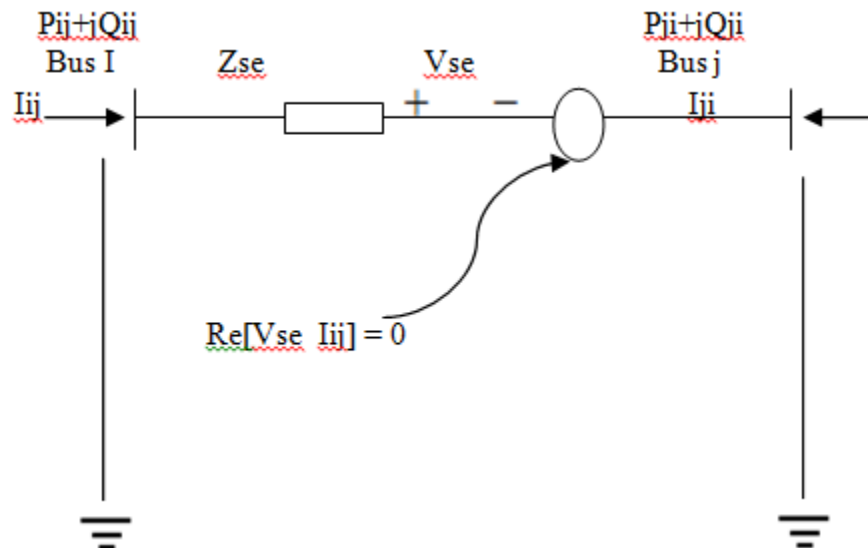


Fig. 2 SSSC Equivalent circuit

IV. OPF FORMULATION WITH SSSC LAGRANGIAN FUNCTION

Based on the above equation the Lagrangian function for the SSSC may be written as,

$$L(x, \lambda) = f(P_g) + \lambda' h(P_g, V, \theta, \delta_{cr}, V_{cr}) \quad (20)$$

In this expression, $f(P_g)$ is the objective function to be optimized $h(P_g, V, \theta, V_{CR}, \delta_{CR})$; represents the power flow equations; x is the vector of state variables; λ is the vector of Lagrange multipliers for

equality constraints; and P_g, V and θ are the active power generation, voltage magnitude, and voltage phase angle, respectively. The SSSC control variables are δ_{CR} and V_{CR} . The inequality constraints $g(P_g, V, \theta, V_{CR}, \delta_{CR}) < 0$ are not shown in Eqn. (3.81) because it is added only to $L(x, \lambda)$ when there are variables outside limits.[19]

The Lagrangian function, $L_{km}(x, \lambda)$ corresponding to the power flow mismatch equations at buses k and m , is given by the following equation,

$$L_{km}(x, \lambda) = \lambda_{pk}(P_k + P_{dk} - P_{gk}) + \lambda_{qk}(Q_k + Q_{dk} - Q_{gk}) + \lambda_{pm}(P_m + P_{dm} - P_{gm}) + \lambda_{qm}(Q_m + Q_{dm} - Q_{gm}) \quad (21)$$

In this expression P_{dk}, P_{dm}, Q_{dk} and Q_{dm} are the active and reactive power loads at buses k and m ; P_{gk}, P_{gm}, Q_{gk} and Q_{gm} are the scheduled active power generations at buses k and m ; and $\lambda_{pk}, \lambda_{pm}, \lambda_{qk}$ and λ_{qm} are Lagrange multipliers at buses k and m . The vector of state variables x is $[V \ \delta]^T$, where V and λ include both nodal voltages and SSSC voltage sources.

4.1 SSSC Power Flow Constraints

The power injected at bus m by the SSSC can be formulated as a power flow constraint in the branch connecting buses m and l . We may write

$$L_{ml}(x, \lambda) = \lambda_{pml}(P_{ml} - P_{specified}) + \lambda_{qml}(Q_{ml} - Q_{specified}) \quad (22)$$

Where, λ_{pml} is the Lagrange multiplier associated with the active power injection at bus m ; and $P_{specified}$ is the specified active power leaving bus m .

In conventional OPF formulations, such constraints are enforced only if power flow limits have been exceeded. However, in this particular application this constraint may remain active throughout the iterative solution [12]. The SSSC Lagrangian function comprising the individual contributions presented above is as follows

$$L_{SSSC}(x, \lambda) = L_{km}(x, \lambda) + L_{se}(x, \lambda) + L_{ml}(x, \lambda) \quad (23)$$

4.2 Optimal setting of SSSC Parameters

Basically, the SSSC has one voltage source inverter (VSI) having a dc storage capacitor. It is connected to the system through a coupling transformer. In this study, the series compensation $\Delta U_{FACTS} = \Delta U_{SSSC}$ is employed [12].

The injected currents at bus i and bus j can be expressed as follows,

$$\Delta I_{i,send} = \frac{\Delta U_{SSSC}}{Z_{ij}} \quad (24)$$

$$\Delta I_{j,send} = -\frac{\Delta U_{SSSC}}{Z_{ij}} \quad (25)$$

The cost function for SSSC can be expressed as follows,

$$C_{SSSC} = 0.0015s^2 - 0.5130s + 133.15 \quad \$/hr \quad (26)$$

where, C_{SSSC} is the cost function of the SSSC in \$/hr and 's' is the operating range of the FACTS devices in MVar.

V. SEVERITY OF OVER LOADABILITY INDEX (SOL) COMPUTATION

The location of the FACTS devices in this work is decided based on the severity of the overloading of that particular branch in which the device is incorporated. The process of ranking the branches based on their load ability in the order of their severity involves the following steps.

Step1: Establish the criterion to be considered in formulating the ranking

Step2: For the criterion established in (Step 1), define a scalar mathematical function which has a

large value of branch load that which stress the system relative to that criterion, and a small value for those which do not; this function is called a "SOL index."

The SOL index is such that contingencies resulting in system conditions yielding large valued over load indices are considered more severe than system conditions with smaller over load indices [16]. In the overload ranker, the SOL index is defined as,

$$SOL = \sum_{i=1}^n \left(\frac{P_i}{P_{i,max}} \right)^2 \quad (27)$$

where,

P_i is the real power flow in line "i",

$P_{i,max}$ is the maximum of active power transfer over the i^{th} line and

'n' is the set of monitored lines contributing to SOL.

5.1 Calculation of SOL for IEEE 5 Bus system

Table 1: SOL index of all buses by running the general OPF for IEEE 5 bus system

Bus No./Node No.	SOL index of each bus	Ranking
[3]	0.5812	1
[4]	0.5310	2
[5]	0.3285	3

As compared the above SOL-indices for the IEEE 5 bus system among the 3 load buses (3, 4, 5) the bus 3 is having the maximum SOL index, it is considered to be the critical bus. Hence line indices will provide accurate information with regard to the stability condition of the lines.

5.2 Calculation of SOL for IEEE 30 Bus system

Table 2: SOL-indices by running the general OPF of maximum loaded buses in IEEE 30 bus system

Branch Number	SOL indices of different branches	Ranking
[30]	0.7776	2
[24]	0.5672	4
[29]	0.8873	1
[28]	0.7486	3
[26]	0.5491	5

As we considered the SOL-index table of the IEEE 30 bus system there will be the 5 load buses (24, 26, 28, 29, 30) with the bus (29) is having the maximum load ability, it is considered to be the critical bus. The branch connected to that particular weakest or critical bus will be the optimal location for the FACTS device to be placed. Hence the branch [13]-[14] is chosen to be the optimal location in the IEEE 30 bus case.

VI. SIMULATION RESULTS

The proposed methodology of active power optimal power flow of series FACTS devices as TCPST, TCSC and SSSC for transmission network is implemented using MATLAB on the two test systems viz., IEEE 5-bus and IEEE-30-bus test systems. The cost constants and their typical values for this problem are energy cost ($a=60$ Rs/hr, $b=200$ Rs/MW/hr, $c=140$ Rs/ MW^2 /hr). The payback period assumed is five years with depreciation factor (α) as 0.1.

6.1 Optimal power flow for IEEE 5 bus system without and with FACTS devices

The 5-bus test system used to illustrate the use of the conventional power flow Newton– Raphson method and is also used to illustrate the use of the OPF and associated data. The maximum and minimum voltage magnitude limits at all buses are taken to be 0.9p.u. and 1.1 p.u., respectively, except at bus 1 (Slack bus), where the maximum limit is set at 1.5 p.u.

The cost coefficients of the two generating buses are taken to be

$$a = 2700 \text{ Rs} / Mw^2 / hr$$

$$b = 153 \text{ Rs} / Mw / hr \quad \text{and,}$$

$$c = 0.18 \text{ Rs} / hr .$$

The maximum and minimum generator active power limits are set at 200MW and 10MW, respectively, whereas the maximum and minimum reactive power limits are set at 300MVAR and -300MVAR, respectively.

Table 3 Comparison of OPF solution without and with FACTS devices for IEEE 5 bus system

Without device	TCPST	TCSC	SSSC
V_{bus} p.u.	V_{bus} p.u.	V_{bus} p.u.	V_{bus} p.u.
1.060	1.06	1.060	1.060
1.100	1.012	1.035	1.049
1.078	1.028	1.042	1.048
1.077	1.027	1.028	1.036
1.072	1.020	1.032	1.049
-----	1.009	1.018	1.027

Table 4 Comparison of Active power and Active power generation cost of different cases for IEEE 5 bus system

Case type	P_{loss} (MW)	$P_{Gen, cost}$ (Rs/hr)
Without device	3.55	34,046
TCPST	3.45	33,874
TCSC	3.41	33,558
SSSC	3.21	33,366

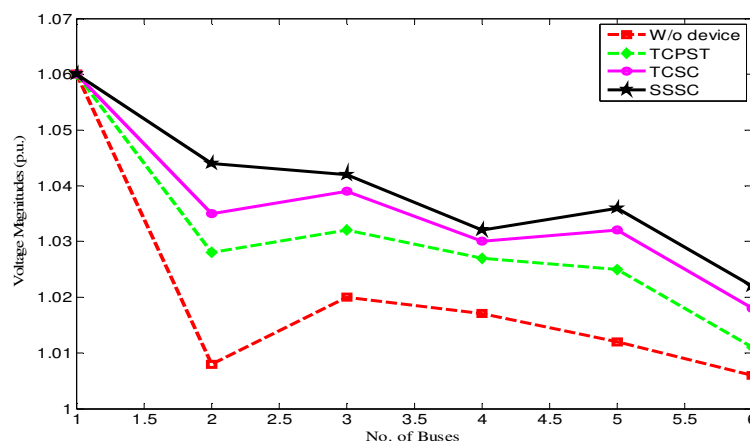


Fig.3 Comparison of without devices and with TCPST, TCSC and SSSC voltages magnitudes at each bus in IEEE 5 Bus system

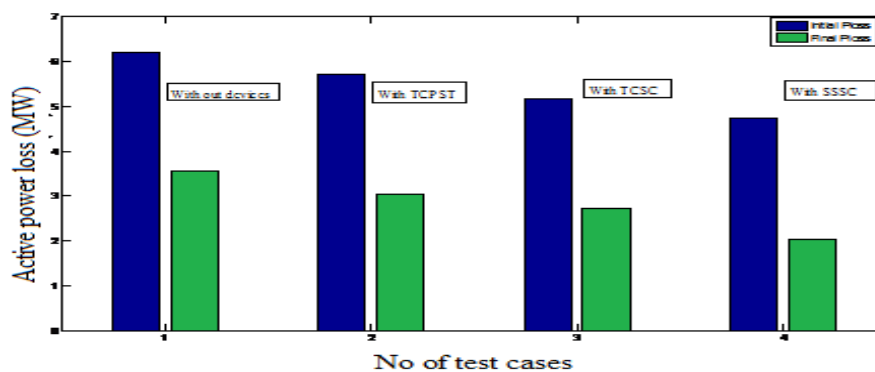


Fig.4 Comparison of Active power loss and % of P_{loss} without and with FACTS devices

6.2 Optimal power flow for IEEE 30 bus system without and with FACTS devices

The IEEE 30-bus test system is used to study the impact of the FACTS devices on the network. The optimal location will be analyzed by using the SOL index from the Table 2 and then FACTS devices are added in series with transmission line (28–29), and the dummy bus (31) is added to enable such a connection to take place.

As compared the voltage magnitudes without and with FACTS devices the voltage profile of the SSSC is improved slightly when compared with TCPST and TCSC, for both the IEEE Test systems. By comparing the SOL-index under normal situation the optimal location of the FACTS device is decided. Hence for the IEEE 30 Bus system (28-29) is the optimally decided branches for the FACTS devices to be incorporated in the electrical power system.

Table 5 The active power and reactive power of different buses for IEEE 30 bus system

Case type	Active power loss (MW)	Reactive power loss (MVar)
Without Device	18.58	52.73
With TCPST	18.49	46.41
With TCSC	18.37	49.23
With SSSC	18.02	42.45

Table 6 The initial and final costs of active power generation at different buses

Case type	IEEE 5 bus system		IEEE 30 bus system	
	$P_{Gen, initial cost}$ (Rs/hr)	$P_{Gen, final cost}$ (Rs/hr)	$P_{Gen, initial cost}$ (Rs/hr)	$P_{Gen, final cost}$ (Rs/hr)
Without Device	35,000	34,046	36,900	36,765
With TCPST	35,000	33,874	36,900	35,325
With TCSC	35,000	33,558	36,900	33,795
With SSSC	35,000	33,366	36,900	33,075

Furthermore, with SSSC the generation cost is reduced to 680 Rs /hr in 5 bus system and 3825 Rs/hr in 30 bus system respectively when compared to TCPST, TCSC and with the base case i.e. 4.66% reduction in the active power generation cost compared to (2.76%, 3.25% and 4.12% for the without FACTS, with TCPST and TCSC respectively).

Table 7 Comparison of Total Active power generation costs in both the test cases

Case Type device	IEEE 5 bus system		IEEE 30 bus system	
	$P_{loss, Total}$ (MW)	$P_{Gen, cost, Total}$ (Rs/hr)	$P_{loss, Total}$ (MW)	$P_{Gen, cost, Total}$ (Rs/hr)
Without device	3.55	34,046	18.58	36,765
TCPST	3.45	33,874	18.49	35,325
TCSC	3.41	33,558	18.37	33,795
SSSC	3.21	33,366	18.02	33,075

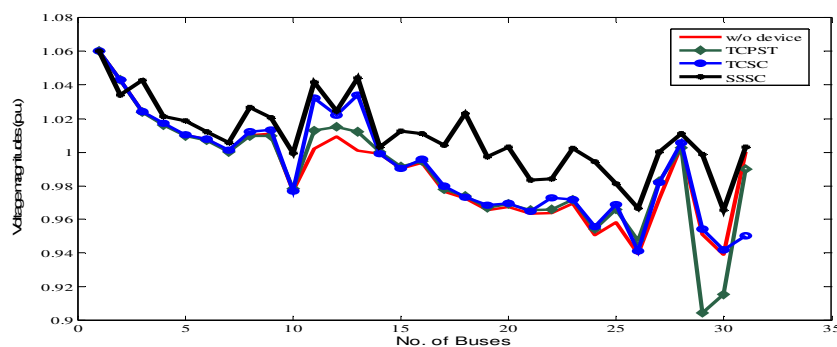


Fig. 5 Comparison of without and with TCPST, TCSC and SSSC voltages at different buses in IEEE 30 Bus system

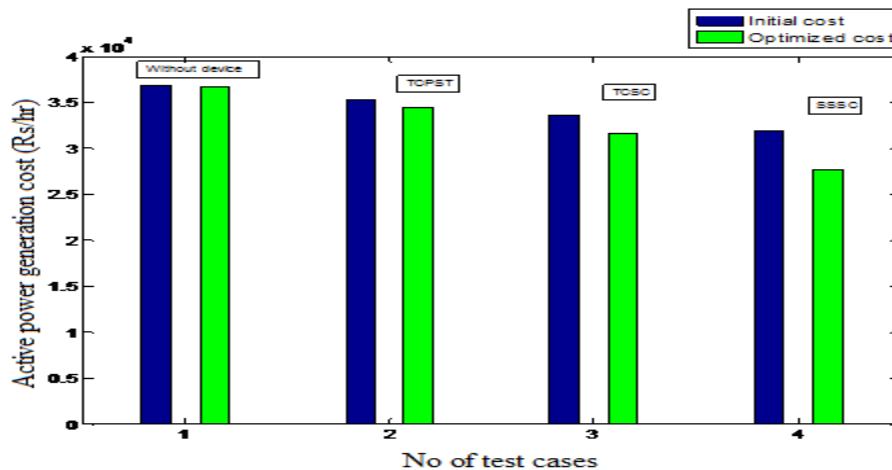


Fig. 6 Comparison of Active power generation cost (Pgi, cost) in case of without device & with TCPST, TCSC and SSSC at different buses in IEEE 30 Bus system

Table 8 Comparison of the costs of different FACTS devices in IEEE 5 bus & IEEE 30 bus system

Type of Device	Cost of the FACTS devices [(Rs/hr) for TCSC & SSSC & (Rs) for TCPST]	
	IEEE 5 bus system	IEEE 30 bus system
TCPST	1,86,075 Rs	3,66,075 Rs
TCSC	6,938.99 Rs/hr	17,600 Rs/hr
SSSC	5,952.47 Rs/hr	13,452 Rs/hr

From Tables (3) and (6) it was clear that the SSSC is improving the overall system performance in both the IEEE test systems when compared with the TCPST, TCSC and base cases. Being having the separate DC link for the SSSC it is capable of generating the reactive power without any external source and hence capable of controlling the reactive power flow of the system like UPFC, instead of controlling only active power as TCSC. The Table 5.18 will gives the costs of the FACTS devices among which the SSSC is possessing the less cost comparable to the other devices. Hence it was concluded that the SSSC is more suitable for the power flow control and optimizing techniques.

VII. CONCLUSIONS

Simulation studies using MATLAB programming code, on IEEE 5 and IEEE 30 bus system are presented to illustrate the methodology and to demonstrate the benefit of the proposed approach. In this paper the optimal power flow analysis with the inclusion of TCPST, TCSC and SSSC has been done. Newton-Raphson method used in polar co-ordinate form is effectively applied to solve the Optimal power flow equations of IEEE 5-Bus & IEEE 30-Bus systems which differ from each other in size & degree of operational complexity.

The modeling of Series FACTS devices are incorporated into an existing Newton-Raphson load flow algorithm, which is capable of solving large power networks very reliably. It shows that Series FACTS environment can be set to control active power, active power loss and voltage magnitude simultaneously. The Severity of Overloading index is used to find location of the series FACTS devices that they are used in the network.

The proposed algorithms were implemented to find out the proper setting and installation cost of the TCPST, TCSC and SSSC in IEEE-5 bus & IEEE-30 bus test systems. By comparing the results, it is observed that SSSC is more effective than TCSC and TCPST in terms of voltage regulation, power loss reduction, minimization of active power generation cost and improving the active power flow.

ACKNOWLEDGMENT

There are several people we would like to thank. First, we would like to thank Dr. Kancharla Ramaiah, correspondent and secretary of Prakasam Engineering College, Kandukur, India for his encouragement and support and providing us with the facility for completing this paper.

REFERENCES

- [1] E.Uzunovic (2001), *EMTP Transient Stability and Power Flow Models and Controls of VSC Based FACTS Controllers*, P.hD Thesis, University of Waterloo, Waterloo, Ontario, Canada
- [2] TjingTlie and wanhong Deng, "Optimal flexible AC Transmission systems devices allocation", *Electric power and energy systems*, Vol. 19, No.2, PP. 125-134, 1997
- [3] J.Bian, D.G.Ramey, R.J.Nelson, and A.Edris (1997), "A Study of Equipment Sizes and Constraints for a Unified Power Flow Controller", *IEEE Trans. On Power Delivery*, Vol.12, No.3, pp. 1385-1391, July.
- [4] N.Li, Y.Xu, and H.Chen (2000), "FACTS Based Power Flow Control in Interconnected Power Systems", *IEEE Trans. On Power Systems*, Vol.15, No.1, pp. 257-262, Feb.
- [5] B.Fardanesh (2004), "Optimal Utilization, Sizing, and steady-State Performance Comparison of Multiconverter VSC-Based FACTS Controllers", *IEEE Trans. On Power Delivery*, Vol.19, No.3, pp. 1321- 1327, July.
- [6] Squives R.B. 1961. Economic dispatch of generation directly from power system voltage and admittances, *IEEE Trans. On Pas*-79(3): 1235-1244.
- [7] S.N. Singh, A.K. David, "Optimal location of FACTS devices for congestion management", *Electric power system research* 58 (2001) 71-79.
- [8] Gyugyi L, Shauder CD, Sen KK (1997) Static synchronous series compensator: a solid- state approach to the series compensation of transmission lines", *IEEE Transactions on Power Delivery*; vol 12, no 1, pp 406-413.
- [9] Sen KK (1998) SSSC - Static synchronous series compensator: theory, modeling, and applications. *IEEE Transactions on Power Delivery*, vol. 13, no 1, pp 241-246
- [10] F.A.L. Jowder. "Influence of mode of operation of the SSSC on the small disturbance and transient stability of a radial power system", *IEEE Transactions on Power Systems*, vol. 20, no 2, May 2005, pp: 935-942.
- [11] Xiao-Ping Zhang, "Advanced modeling of the multi control functional Static Synchronous series compensator (SSSC) in Newton Power flow", *IEEE Transactions on Power Systems*, vol. 18, no. 14, November 2003, pp. 1410-1416.
- [12] X. P. Zhang. "Modeling of FACTS in power flow and optimal power flow analysis", *DianliXitongZidonghua/Automation of Electric Power Systems*, vol.29, no.16, Aug-25, 2005, pp: 22-29+65.
- [13] C. R. Fuerte-Esquivel and E. Acha, "A Newton-type algorithm for the control of power flow in electrical power networks," *IEEE Trans. Power Systems*, vol.12, no.4, pp.1474-1480, Nov. 1997.
- [14] C. R. Fuerte-Esquivel, E. Acha, S. G. Tan and J. 1. Rico, "Efficient object oriented power systems software for the analysis of large-scale networks containing FACTS-controlled branches," *IEEE Trans. Power Systems*, vol.13, no.2, pp.464-472, May, 1998.
- [15] D. Devaraj, R. NarmathaBanu "Optimal Power Flow for Steady State Security Enhancement using Enhanced Genetic Algorithm with FACTS Devices" *Asian Power Electronics Journal*, Vol. 4, No.3 December 2010
- [16] MarayatiMarsadek, Azah Mohamed, ZulkifiMohd. Norpiah "Assessment and classification of line overload risk in power systems considering different types of severity functions" *WSEAS transactions on power systems*, Issue 3, Volume 5, July 2010
- [17] K.ChandrasekaranK.AruljeyarajL.SahayasenthamilDr.M.saravanan "New Method To Incorporate Facts Devices In Optimal Power Flow Using Particle Swarm Optimization" *Journal of Theoretical and Applied Information Technology*
- [18] A. B.Bhattacharyya, B. S.K.Goswami "OPTIMAL Placement of FACTS Devices by Genetic Algorithm for the Increased Load Ability of a Power System" *World Academy of Science, Engineering and Technology* 75 2011
- [19] H O Bansal, H P Agrawal, S Tiwana, A R Singal And L Shrivastava "Optimal Location of FACT Devices to Control Reactive Power" H.O. Bansal et al. / *International Journal of Engineering Science and Technology* Vol. 2(6), 2010, 1556-1560

- [20] K. Ravi , M. Rajaram, J. Belwin Edward, "Hybrid Particle Swarm Optimization Technique for Optimal Location of FACTS devices using Optimal Power Flow" European Journal of Scientific Research ISSN 1450-216X Vol.53 No.1 (2011), pp.142-153

Author Profiles

M. Balasubba Reddy received the Bachelor of Engineering degree in Electrical and Electronics Engineering from Madras University in 2000 and Master's degree from NIT Trichy in 2004. He is a research student of JNTU Kakinada. Currently, he is a Professor and head of Dept. of EEE at Prakasam Engineering College, Kandukur. His areas of interests are in Power system, power electronics and FACTS controllers.



Y. P. Obulesu received his B.E degree in Electrical Engineering from Andhra University, Visakhapatnam in 1996. M.Tech degree from IIT, Kharagpur, in 1998. He received his PhD degree from Jawaharlal Nehru Technological University, Hyderabad, in 2006. Currently he is working as a Professor and head of Dept. of EEE at LBRCEC, Mylavaram. He has published several National and International Journals and Conferences. His area of interest is the simulation and design of power electronics systems, DSP controllers.



S. Sivanagaraju received his Master's degree in 2000 from IIT, Kharagpur and did his Ph.D from J.N.T. University in 2004. He is currently working as associate professor in the department of Electrical Engineering. J.N.T.U. College of Engg. (Autonomous) Anantapur, Andhra Pradesh, India. He is referee for IEE Proceedings-Generation Transmission and Distribution and International journal of Emerging Electrical Power System. He has 40 publications in National and International journals and conference to his credit. His areas of interest are in Distribution Automation, Genetic Algorithm application to distribution systems and power system.



MODELING AND SIMULATION OF THE PATCH ANTENNA BY USING A BOND GRAPH APPROACH

Riadh Mehouchi¹, Hichem Taghouti¹, Sameh Khmailia¹ and Abdelkader Mami^{1,2}

¹Department of Electrical Engineering, Laboratory of Analysis and Command Systems,
National Engineering School of Tunis, P.O. Box 37, 1002 Tunisia

²Laboratory of High Frequency Circuits, Department of Physics, Sciences Faculty of Tunis,
2092 El Manar Tunis Tunisia

ABSTRACT

Further to our own studies carried out previously on the application of the bond graph approach on antennas based on localized elements, we tried in the continuation of our research to devote our studies on the antennas based on distributed elements. For that purpose, we have chosen as first departure, a patch antenna of which we want to find its scattering parameters by our own and new methodology. So we can say that in this paper, we propose a new methodology to study a patch antenna by applying the bond graph approach and the scattering formalism. This study permits us on the one hand to determine and to simulate the scattering parameters (reflexion and transmission coefficients S_{11} , S_{12} , S_{21} and S_{22}) of a patch antenna and on the other hand to modelling the incident and reflected wave propagation for all the points of this antenna by the bond graph approach.

KEYWORDS: Patch Antenna, Scattering Matrix, Electrical Model, Microstrip Line, Scattering Parameters, Scattering Bond Graph & Simulation.

I. INTRODUCTION

The microstrip patch antenna [4] is used in many applications in the communication systems [3] because of their ease of fabrication, low-profile, low cost and small size. In the setting of this article, we look to conceive and to use a simple and precise electrical model of square patch keeping in mind all the electric and geometric characteristics of elements radiating and of their feedings [3,4]. For that we will use our new methodology called “Scattering Bond Graph Modelling” [11,12] which combine on the same time the Scattering Formalism [11] and the Bond Graph Approach [21] which is a graphic language unified for all the fields of the engineering and confirmed like a structured approach with the modelling and simulation of many systems. The purpose of this paper is to present and apply this new methodology to extract the scattering parameters of our studied system (patch antenna) while basing on its causal and reduced bond graph model [21].

At first, and after having to determine and simulate the scattering matrix of the studied antenna from its electrical model, we propose to use the causal bond graph model of this patch antenna [13] to find, on one hand, the integro-differentials operators [6] which is based on the causal ways and loops present in the bond graph model and, on the other hand, to extract the wave matrix [8] from these operators.

Then, we extract directly the scattering parameters [11] from the found wave matrix [8] and, at the aim to validate the found results; we make a comparison by the simulation of these scattering parameters with a simple program and the classic techniques of conception and simulation of the microwave circuits [20].

Finally, we will propose in our future work, to build a particular type of bond graph model which is able to highlight these transmission and reflection coefficients (Scattering parameters) [1].

II. DETERMINATION OF THE SCATTERING MATRIX FROM THE PATCH ANTENNA

At this part, we will determine the reflected and the transmission coefficients (S_{11} , S_{12} , S_{21} and S_{22}) of this proposed antenna [3,19] after a small explication of this patch and by a joint use of the scattering formalism and the bond graph approach and so we will indicate the biggest application of this new methodology (Scattering bond graph).

2.1. Design of the Square Antenna

The figure 1 represents the geometry of a square micro-strip patch on a dielectric substrate with a ground plane. The antenna is mounted on a substrate material with a thickness of $H = 3.2mm$ and it has an edge of $W = 36mm$, a relative permittivity: $\epsilon_r = 2.6$ and loss tangent: $\tan(\delta) = 0.002$. The dimension 'W' of square edge is calculated using the following equation [19].

$$W = \frac{1}{2f_r} \frac{c}{\sqrt{\epsilon_r}} \quad (1)$$

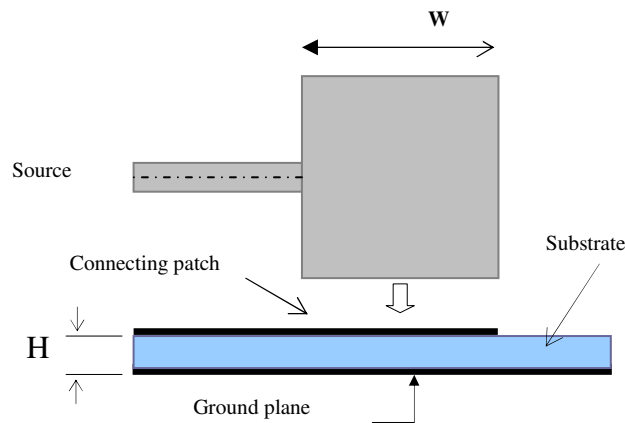


Figure1. Square patch antenna

2.2. Determination of the parameters of the antenna

At first, we are going to calculate the frequency of resonance f_r , and the total quality factor Q_T , in following we will deduct the other parameters from the following equation [7, 9].

$$C = \frac{(\epsilon_{eff} \epsilon_0 W^2)}{2H} \cos^{-2} \left(\frac{\pi x_0}{W} \right) \quad (2)$$

C: capacitance

x_0 : The distance of the feed point from the edge of the patch.

W: length of the square

H: thickness of dielectric

The resonant resistance R is calculate using the following equations.

$$R = \frac{Q_T}{w_r C} \quad (3)$$

$$Q_T = \left(\frac{1}{Q_R} + \frac{1}{Q_C} + \frac{1}{Q_D} \right)^{-1} \quad (4)$$

Q_R : Radiation quality factor.

Q_C : Losses in the conductor.

Q_D : Losses in the dielectric.

$$Q_R = \frac{c_0 \sqrt{\epsilon_{dyn}}}{4f_r H} \quad (5)$$

$$Q_D = \frac{1}{\tan(\delta)} \quad (6)$$

$$Q_C = \frac{0.786 \sqrt{f_r Z_{a0}(W) H}}{P_a} \quad (7)$$

$$Z_a(W) = \frac{60\pi}{\sqrt{\epsilon_r}} \left\{ \frac{W}{2H} + 0.441 + 0.082 \left(\frac{\epsilon_r - 1}{\epsilon_r^2} \right) + \frac{(\epsilon_r + 1)}{2\pi\epsilon_r} \left[1.451 + \ln \left(\frac{W}{2H} + 0.94 \right) \right] \right\}^{-1} \quad (8)$$

$$Z_{a0}(W) = Z_a(W, \epsilon_r = 1) \quad (9)$$

Z_a : Is the impedance of an air filled microstrip line.

$$P_a(W) = \frac{2\pi \left(\frac{W}{H} + \frac{\frac{W}{H\pi}}{\frac{W}{2H} + 0.94} \right) \left(1 + \frac{W}{H} \right)}{\left\{ \frac{W}{H} + \frac{2}{\pi} \left[2\pi \exp \left(\frac{W}{2H} + 0.94 \right) \right] \right\}^2} \quad (10)$$

$\tan(\delta)$: is the tangent of loss in the dielectric and it given by the following equation.

$$\delta = \left[\frac{H}{W} 0.882 + \frac{0.164(\epsilon_r - 1)}{\epsilon_r^2} + \frac{(\epsilon_r + 1) \left(0.756 + \ln \left(\frac{W}{H} + 1.88 \right) \right)}{\pi\epsilon_r} \right] \quad (11)$$

$$\epsilon_{dyn} = \frac{C_{dyn}(\epsilon)}{C_{dyn}(\epsilon_0)} \quad (12)$$

$$C_{dyn}(\epsilon) = \frac{\epsilon_0 \epsilon_r A}{H \gamma_n \gamma_m} + \frac{1}{2\gamma_n} \left(\frac{(\epsilon_{reff}(\epsilon_r, H, W))}{c_0 Z(\epsilon_r = 1, H, W)} - \frac{\epsilon_0 \epsilon_r A}{H} \right) \quad (13)$$

$$\gamma_j = \begin{cases} 1, j = 0 \\ 2, j \neq 0 \end{cases} \quad (14)$$

$$Z(W, H, \epsilon_r) = \frac{377}{H} \left[\frac{W}{H} + 1.393 + 0.667 \ln \left(\frac{W}{H} + 1.44 \right) \right]^{-1} \quad (15)$$

$$f \left(\frac{W}{H} \right) = 6 + (2\pi - 6) \exp \left(-\frac{30.66}{\frac{W}{H}} \right) 0.7528 \quad (16)$$

The effective dielectric constant of a microstrip line is given bellow [10]:

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + 12 \frac{H}{W} \right]^{-\frac{1}{2}} \quad (17)$$

The inductance L is given by the following equation:

$$L = \frac{1}{(w_{res})^2 C} \quad (18)$$

$$w_{res} = 2\pi f_r \quad (19)$$

2.3. Electrical model of the square patch antenna

The structure can be modelled as an RLC resonant circuit [7]. The model parameters are calculated using the formulas developed in section 1. The patch is excited by a transmission line and the dimension is determined to resonate frequency equal 2.45GHz.

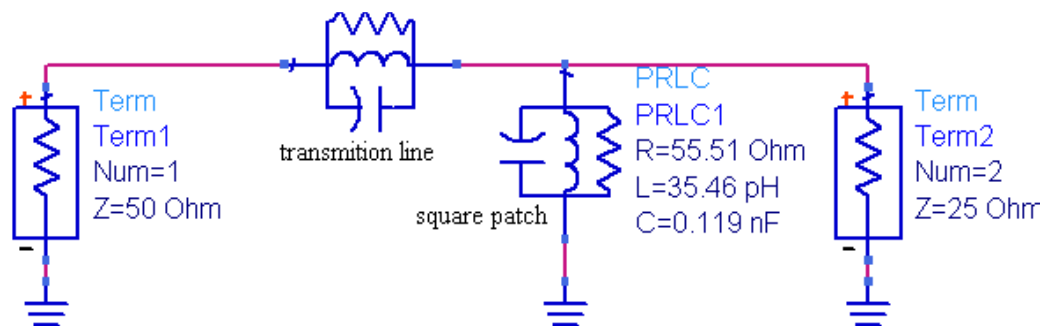


Figure2. Electrical model of the square patch antenna

2.4. Simulation results of scattering parameters

The tradition tools for simulation of the scattering parameters under HD-ADS software [13] as figure 3 refer to, visualize the coefficients of reflection.

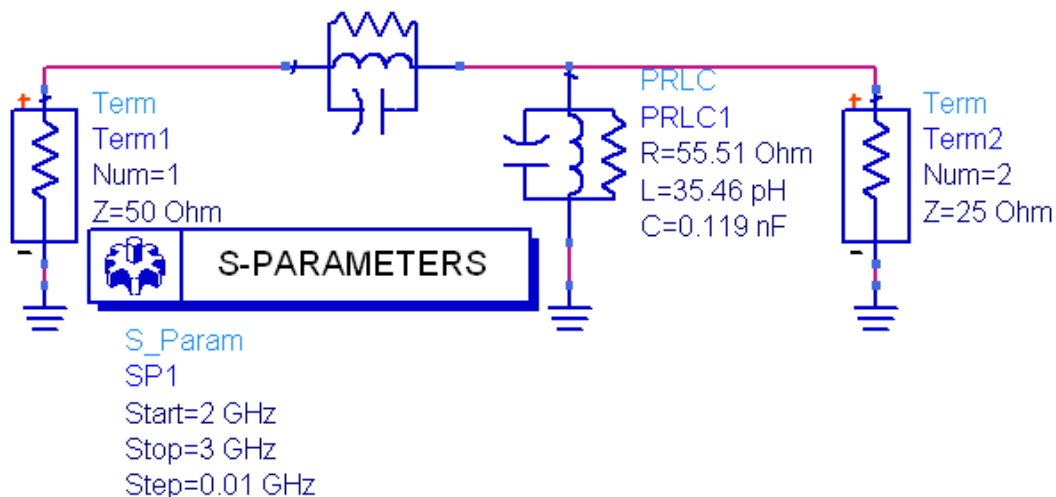


Figure3. Square patch antenna, under HP-ADS

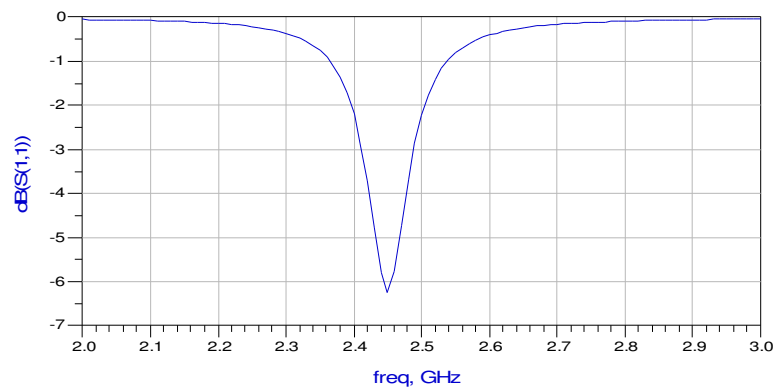


Figure4. Reflection coefficient (S_{11}) simulated of square patch antenna under HP-ADS

III. SCATTERING PARAMETERS OF THE SQUARE PATCH ANTENNA EXPLOITED FROM ITS BOND GRAPH MODEL

In this part we propose to determine the scattering parameters [12] of the patch antenna from its transformed and reduced bond graph model [13] without forgetting to pay attention to causality assignment.

3.1. Relation between bond graph model and wave scattering matrix

We can represent every physical system by a quadripole inserted between two ports P_1 and P_2 who respectively represent the source and the load of all the system [5]. This system can be represented by a generalized bond graph model transformed and reduced as figure 5 indicates it [17, 18].



Figure5. General transformed and reduced bond graph model

- ϵ_1 and ϵ_2 are respectively the reduced variable (effort) at the entry and the exit of the system.
 - ϕ_1 and ϕ_2 are respectively the reduced variable (flow) at the entry and the exit of the system.
- The causality assignment to the reduced bond graph model of figure 5 enables us to notice that there are four different cases of causality assignment in input-output of the process [14, 15].

3.1.1. Case 1: Flow-Effort Causality



Figure6. Reduced bond graph model with flow-effort causality

From this type of assignment causality we can deduce the following matrix.

$$\begin{bmatrix} \varepsilon_1 \\ \varphi_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varepsilon_2 \end{bmatrix} \quad (20)$$

3.1.2. Case 2: Effort-Flow Causality



Figure7. Reduced bond graph model with effort -flow causality

From this type of assignment causality we can deduce the following matrix.

$$\begin{bmatrix} \varphi_1 \\ \varepsilon_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varphi_2 \end{bmatrix} \quad (21)$$

3.1.3. Case 3: Flow-Flow Causality



Figure8. Reduced bond graph model with flow -flow causality

From this type of assignment causality we can deduce the following matrix.

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} \quad (22)$$

3.1.4. Case4: Effort-Effort Causality



Figure9. Reduced bond graph model with effort -effort causality

From this type of assignment causality we can deduce the following matrix.

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (23)$$

We can note that

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \quad (24)$$

H_{ij} represent the integro-differentials operators associated to the causal ways connecting the port P_j to the port P_i and obtained by the general form given below [16, 17].

$$H_{ij} = \sum_{k=1}^n \frac{T_k \Delta_k}{\Delta} \quad (25)$$

$$\Delta = 1 - \sum B_i + \sum B_i B_j - \sum B_i B_j B_k + \dots + (-1)^m \sum \dots \quad (26)$$

- Δ = the determinant of the causal bond graph
- H_{ij} = complete gain between P_j and P_i .
- P_i = input port.
- P_j = output port.
- n = total number of forward path between P_i and P_j
- T_k = gain of the k^{th} forward path between P_i and P_j
- B_i = loop gain of each causal algebraic loop in the bond graph model.
- $B_i B_j$ = product of loop gains of any two non-touching loops (no common causal bond).
- $B_i B_j B_k$ = product of the loop gains of any three pairwise nontouching loops.
- Δ_k = the factor value of Δ for the k^{th} forward path, this value calculates himself as Δ when one only keeps the causal loops without touching the k^{th} chain of action.

We noted that:

$$a_i = \frac{\varepsilon_i + \varphi_i}{2}, \quad a_i = \frac{\varepsilon_i - \varphi_i}{2} \quad (27)$$

$$\varepsilon_i = \frac{V}{\sqrt{R_0}}, \quad \varphi_i = I\sqrt{R_0} \quad (28)$$

These are reduced voltage and current.

$$\begin{pmatrix} \varepsilon_1 \\ \varphi_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad (29)$$

$$\begin{pmatrix} \varepsilon_2 \\ -\varphi_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \quad (30)$$

$$\begin{pmatrix} b_1 \\ a_1 \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = (W) \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \quad (31)$$

By using the preceding equations; we can find for each case of causality one wave matrix.

$$W = \frac{1}{2H_{21}} \begin{bmatrix} 1-H_{11}+H_{22}-\Delta H & -1+H_{11}-H_{22}-\Delta H \\ -1-H_{11}-H_{22}-\Delta H & 1+H_{11}-H_{22}-\Delta H \end{bmatrix} \quad (32)$$

$$W = \frac{1}{2H_{21}} \begin{bmatrix} 1-H_{11}+H_{22}-\Delta H & 1-H_{11}-H_{22}+\Delta H \\ 1+H_{11}+H_{22}+\Delta H & 1+H_{11}-H_{22}-\Delta H \end{bmatrix} \quad (33)$$

$$W = \frac{1}{2H_{21}} \begin{bmatrix} 1-H_{11}+H_{22}-\Delta H & -1+H_{11}+H_{22}-\Delta H \\ 1+H_{11}+H_{22}+\Delta H & 1+H_{11}-H_{22}-\Delta H \end{bmatrix} \quad (34)$$

$$W = \frac{1}{2H_{21}} \begin{bmatrix} -1+H_{11}-H_{22}+\Delta H & 1-H_{11}-H_{22}+\Delta H \\ -1-H_{11}-H_{22}-\Delta H & 1-H_{11}+H_{22}-\Delta H \end{bmatrix} \quad (35)$$

With:

$$\Delta H = H_{11}H_{22} - H_{12}H_{21} \quad (36)$$

The following scattering matrix gives us the scattering parameters [15, 16] [17]:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (S) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (37)$$

The relation between equations and gives us the following equations:

$$\begin{cases} w_{11} = -s_{22} s_{21}^{-1} \\ w_{12} = s_{21}^{-1} \\ w_{21} = (s_{12}s_{21} - s_{11}s_{22}) s_{21}^{-1} \\ w_{22} = s_{11} s_{21}^{-1} \end{cases} \quad (38)$$

The corresponding scattering matrix [16, 17] is given by:

$$S^T = \begin{pmatrix} w_{12}w_{22}^{-1} & [w_{11}w_{22} - w_{21}w_{12}]w_{22}^{-1} \\ w_{22}^{-1} & -w_{21}w_{22}^{-1} \end{pmatrix} \quad (39)$$

3.2. Application to the square patch antenna

The bond graph model of the square is giving bellow:

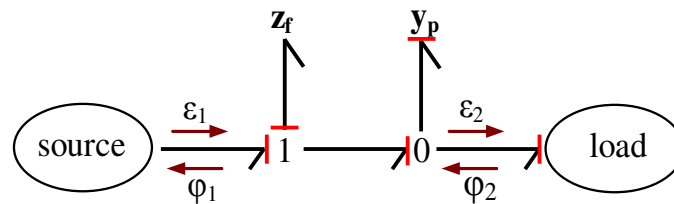


Figure10. Reduced and causal bond graph model of the square patch antenna

- z_f : The reduced equivalent impedance of the feeding line.
- y_p : The reduced equivalent admittance of the patch.

$$z_f = \tau_{C1}s + \frac{1}{\tau_{L1}s} + \frac{1}{\tau_{R1}} \quad (40)$$

$$y_p = \tau_{C2}s + \frac{1}{\tau_{L2}s} + \frac{1}{\tau_{R2}} \quad (41)$$

Where:

$$\tau_{Ci} = R_0 * C_i \quad (42)$$

$$\tau_{Li} = \frac{L_i}{R_0} \quad (43)$$

s : The Laplace operator and R_0 : the scaling resistance.

The sub-model is in conformity with case (2) described previously.

$$L_1 = \frac{-1}{z_f y_p} : \text{Loop gain of the algebraic given by sub-model.}$$

$$\Delta_1 = 1 + \frac{1}{z_f y_p} : \text{Determinant of causal bond graph of the model.}$$

The following equations represent the integro-differentials [17, 18] operators of the model.

$$\left\{ \begin{array}{l} H_{11} = \frac{z_f}{z_f y_p + 1} \\ H_{12} = \frac{1}{z_f y_p + 1} \\ H_{21} = \frac{1}{z_f y_p + 1} \\ H_{22} = \frac{-y_p}{z_f y_p + 1} \\ \Delta H = \frac{-z_f y_p}{z_f y_p + 1} \end{array} \right. \quad (44)$$

From these operators, we can determine the following scattering parameters:

$$S_{11} = \frac{-z_f + y_p + 1}{D(s)} \quad (45)$$

$$S_{12} = S_{21} = \frac{2z_f y_p}{D(s)} \quad (46)$$

$$S_{22} = \frac{-z_f + y_p - 1}{D(s)} \quad (47)$$

$$D(s) = y_p + 2y_p z_f + z_f + 1 \quad (48)$$

3.3. Simulation Results of the Scattering Parameters

To determine the coefficient of reflection, represented by the figure 10 we used the simple programming and simulation of the scattering parameters equations [15, 16].

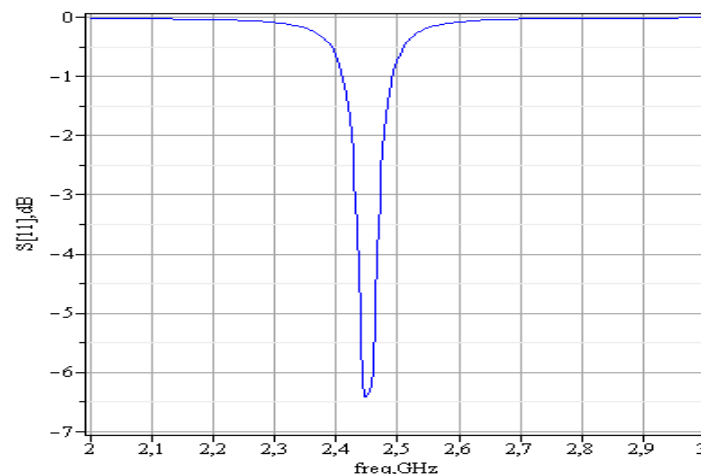


Figure10. Reflection coefficient simulated of square patch antenna using bond graph

The purpose of this simulation is to validate this new extraction method of scattering parameters from a causal bond graph model of a patch antenna we note that it is always necessary taking account of causality concept in order to get the right results as opposed to the work carried out by [13] [15] where the causality concept was ignored.

The simulation given by figure 6 is carried out in the maple software, this figure show the reflection coefficient of the patch antenna that is the same coefficient carried out in the HP-ADS software. We can see that the two results are similar. For the traditional method in microwaves and this new method using bond graph the reflection coefficient have the same resonant frequency and the same gain but with a little difference in Band width because of the difference between calculated and simulated

losses. The figure 6 presents a resonant frequency equal to 2.45 GHz, a gain equal 6.4dB and a bond width about 50MHz, in the figure 4 a bond width about 78MHz [16, 17, 18].

IV. CONCLUSIONS

In this article, we showed a simple method to determinate the scattering parameters of the patch antenna. We validated the results by a simple comparison between the traditional methods used in microwaves under HP-ADS software and simulation by the methods of the reduced and causal bond graph. Now we can apply this technique to a network of antennas and to all microwaves circuits.

ACKNOWLEDGEMENTS

We would like to thank especially *Prof. MAMI Abdelkader* and *Dr. TAGHOUTI Hichem* for the time and guidance given throughout the all carried out works, without forgetting the members of the unit of electronics and high frequency circuits *Mr. MEHOUACHI Riadh* and *Miss. KHMALIA Sameh* and all those who contributed and aided for this study in particularly L.A.C.S members (Laboratory of analysis and command systems).

REFERENCES

- [1] Duclos G, Clément AH (2003). A new method for the calculation of transmission and reflection coefficients for water waves in a small basin. *Comptes Rendus Mécanique*. 331(3): 225-230.
- [2] Ferrero A, Pirola M (2006). "Generalized Mixed-Mode S-Parameters". *IEEE Transactions on Microwave Theory and Techniques*. 54(1): 458-463.
- [3] Hossain, E. Weihua, Z. (2004). "Guest editorial: advances in wireless communications and networking". *Canadian Journal of Electrical and Computer Engineering*, 29(1) : iv – vi.
- [4] Jin, N., Fan Yang, Rahmat-Samii, Y (2006). "A novel patch antenna with switchable slot (PASS): dual-frequency operation with reversed circular polarizations". *IEEE Transactions on Antennas and Propagation*, 54(3) :1031 – 1034.
- [5] Kamel A, Dauphin-Tanguy G (1993). "Bond Graph Modelling of Power Waves in the Scattering Formalism". *J. title: SIMULATION SERIES*. 25(2): 41. Publisher: SCS Society for Computer Simulation, USA, And ISSN: 0735-9276.
- [6] Khachatryan A (2008). "Factorization of a convolution-type integro-differential equation on the positive half line". *60(11): 1555–1567*.
- [7] Kamel A, Dauphin-Tanguy G (1996). "Power transfer in physical systems using the scattering bond graph and a parametric identification approach". *Syst. Anal. Modelling Simulation*, 27(1): 1-13.
- [8] Magnusson PC, Alexander GC, Tripathi VK, Weisshaar A (2001). "Transmission lines and wave propagation". 4th ed. CRC Press.
- [9] Maher A, Scavarda S (1991). "A procedure to match bond graph and scattering formalisms". *J. Franklin Inst.*, 328(5-6): 887-89.
- [10] Molisch AF, Steinbauer M, Toeltsch M, Bonek E, Thomä RS (2002). "Capacity of MIMO Systems Based on Measured Wireless Channels". *IEEE J. on Selected Areas in Communications*, 20(3): 561-569.
- [11] Newton RG (2002). "Scattering theory of waves and particles". New York: Springer-Verlag; Dover Edition.
- [12] Paynter HM, Busch-Vishniac I (1988). "Wave scattering Approaches to conservation and causality". *J. Franklin inst.*, 325(3): 295-313.
- [13] Shamash Y (1980). "Stable biased reduced order models using the Routh method of reduction". *Int. J. Sys. Sci.*, 1464-5319. 11(5): 641- 654.
- [14] Taghouti H, Mami A (2010c). "Modelling Method of a Low-pass Filter based on Microstrip T-Lines with Cut-Off Frequency 10 GHz by the Extraction of its Wave-Scattering Parameters from its Causal Bond Graph Model". *Am. J. Eng. Appl. Sci.*, 3(4): 631-642.
- [15] Taghouti H, Mami A (2009). "Application of the reduced bond graph approaches to determinate the scattering parameters of a high frequency filter". *Proceedings of 10th Int. conference on Sci., and Techniques of Automatic Control & Computer Engineering*. Hammamet, Tunisia, December 20-22, 2009, STA'2009-SSI-548: 379-391.
- [16] Taghouti H, Mami A (2010a). "How to Find Wave-Scattering Parameters from the Causal Bond Graph Model of a High Frequency Filter". *Am. J. Appl. Sci.*, 7(5): 702-710.
- [17] Taghouti H, Mami A (2010b). "Extraction, Modelling and Simulation of the Scattering Matrix of a Chebychev Low-Pass Filter with cut-off frequency 100 MHz from its Causal and Decomposed Bond Graph Model". *ICGST Int. J. Auto. Cont. Sys. Eng.*, 10(1): 29-37.

- [18] Taghouti H, Mami A. "New extraction method of the scattering parameters of a physical system starting from its causal bond graph model: Application to a microwave filter". International Journal of the Physical Sciences. Vol. 6(13), pp. 3016–3030, 4 July, 2011
- [19] Trabelsi H, Gharsallah A, Baudrand H. (2003). "Analysis of microwave circuits including lumped elements based on the iterative method". Int. J. RF Microwave CAE. 13(4): 269-275.
- [20] Vendelin GD, Pavio AM, Rohde UL (2005). "Microwave Circuit Design Using Linear and Non Linear Techniques". 2nd Edition, ISBN: 978-0-471-41479-7, Hardcover, 1080 pages.
- [21] Wilfrid, M., Omar, M., Bogdan, C., Daniel, T., Jérôme, P., Martine, P. (2007). "Bond graph formulation of an optimal control problem for linear time invariant systems". Journal of the Franklin Institute, 345(4) : 349-373.

Authors

RIADH MEHOUCI: Was born in SILIANA (Tunisia) in December 1978. He received the master diploma in electronic (Numeric Analysis and Treatment of Electronics Systems) from the Faculty of Sciences of Tunis (FST) in 2007. Since September 2006, he was been a TECHNICAL TEACHING PROFESSOR in secondary school in Tunis. Actually, he is preparing his thesis at National Engineering School of Tunis (ENIT) in Tunisia. His main research areas are the bond graph for the microstrip patch antenna.



HICHEM TAGHOUTI: Was born in Tunisia in December 1979. He received the master diploma in electronic (Numeric Analysis and Treatment of Electronics Systems) from the Faculty of Sciences of Tunis (FST) in 2005. Since September 2006, he was been an Assistant of Physics-Electronic in Faculty of Sciences of Gafsa in Tunisia. He was preparing his Thesis at National Engineering School of Tunis (ENIT) in Tunisia and faculty of Sciences of Tunis (FST) in Tunisia. His main research areas are the Bond Graph modelling and its New Applications in High Frequencies Domain. On April 2011 he received his PhD. actually he is working as Assistant Professor on Electrical Engineering in Superior institute of sciences and technologies of environment of BORJ-SEDRIA Tunis.



SAMEH KHMALIA: Was born in Gafsa-Tunisia in April 1982. She received the master diploma in electronic (Numeric Analysis and Treatment of Electronics Systems) from the Faculty of Sciences of Tunis (FST) in 2010. From February to September 2010, she was been a contractual assistant of Physics-Electronic in Higher Institute of Accountancy and Entrepreneurial Administration of MANNOUBA in Tunisia. Since September 2010, she was been a contractual assistant of Physics-Electronic in Higher Institute of Applied Sciences and Technology of Gafsa in Tunisia. She prepares her Thesis at National Engineering School of Tunis (ENIT) in Tunisia and faculty of Sciences of Tunis (FST) in Tunisia. Her main research areas are the Scattering- Bond Graph modelling and its New Applications in High Frequencies Domain.



ABDELKADER MAMI: Was born in Tunisia, he is a Professor in Faculty of Sciences of Tunis (FST). He was received his Dissertation H.D.R (Enabling To Direct of Research) from the University of Lille (France) 2003, he is a member of Advise Scientific in Faculty of Science of Tunis (Tunisia), he is a President of commutated thesis of electronics in the Faculty of sciences of Tunis, He is a person in charge for the research group of analyze and command systems in the ACS- laboratory in ENIT of Tunis and in many authors fields.



DESIGN AND VERIFICATION ANALYSIS OF AVALON INTERRUPT INTERFACE WITH COVERAGE REPORT

Mahesh Kumar Jha, Richa Sinha and Akhilesh Kumar
Department of E&C Engineering, NIT Jamshedpur, Jharkhand, India

ABSTRACT

Design is an accurate representation of the specification and is carried out at each step of the manufacturing process. The process of verification parallels the design creation process. Verification shows the correctness of the design, makes sure about no bugs in the design and its functionality. The important part of verification process is coverage analysis; it gives idea that to what degree the source code of the DUT has been tested. Code coverage measures how thoroughly our tests exercised the "implementation" of the design specification, and not the verification plan. Functional Coverage is User specified and not automatic like code coverage. It is based on design specification/intent and not on implementation. Avalon interfaces are used to easily connect components in an FPGA in order to simplify system design. It defines the various Interfaces roles which are used in both high-speed streaming and memory-mapped applications. The Avalon bus is a simple bus architecture which makes the connection between on-chip processors and peripherals together into a SOPC. It is an interface that specifies the port connections between master and slave components, and also specifies the timing by which these components communicate. The Avalon Interrupt Interface is one of the roles of Avalon Interface, which allows the components to signal the events to other components. In this paper, code coverage analysis in design is concluded using QuestaSim simulator and functional coverage analysis in verification is concluded using Riviera simulator..

KEYWORDS: Avalon Interrupt Interface, SystemVerilog, System-on-a-Programmable Chip (SOPC), DUT (Design Under Test), Code Coverage, Functional Coverage, VMM (Verification methodological manual), IP (Intellectual Property), VIP (Verification Intellectual Property)

I. INTRODUCTION

Avalon interface bus is ALTERA's standard interface bus used in SOPC Builder, through which all Avalon compatible cores can easily communicate. SOPC Builder is a powerful system development tool. It enables us to define and generate a complete SOPC in much less time than using traditional, manual integration methods. The Avalon interface family defines interfaces for streaming high-speed data, reading and writing registers and memory, and controlling off-chip devices. These standard interfaces are designed into the components available in Qsys. We can also use these standardized interfaces in our custom components. By using these standard interfaces, we enhance the interoperability of our designs [1]. Avalon Interrupt Interface is one of the interface roles of Avalon Interface which allows components to signal events to other components. The Avalon interface defines a set of signal types, the behavior of these signals, and the types of transfers supported by these signals [2]. In this paper, coverage analysis is done for the design and verification of the Avalon Interrupt Interface.

Coverage analysis is used to measure the effectiveness of verification implementation, a quantitative measurement of the testing space. This analysis states the depth to which the DUT source code has been tested. The goal of verification is to ensure that a design behaves correctly in its real environment [5]. The easiest way to measure verification progress is with code coverage. Here we measure how many lines of code have been executed (line coverage), which paths through the code and expressions have been executed (path coverage), which single-bit variables have had the values 0

or 1 (toggle coverage), and which states and transitions in a state machine have been visited (FSM coverage) [5]. Branch Coverage is one of the types of code coverage which reports the true or false of the conditions like if-else and case statement. Functional coverage is a measure of which design features have been exercised by the tests. Functional coverage is tied to the design intent and is sometimes called “specification coverage,” while code coverage measures the design implementation [5]. In this paper, the verification analysis is done using VMM methodology.

The VMM methodology facilitates the creation of robust, reusable and scalable verification environments using SystemVerilog. It is the most widely used and proven verification methodology for SystemVerilog [8] [9].

II. FEATURES OF AVALON INTERRUPT INTERFACE

Some of the prominent features of the Avalon Interface are [2]:

- *Separate Address, Data and Control Lines* – Provides the simplest interface to on-chip logic. By using dedicated address and data paths, Avalon peripherals do not need to decode data and address cycles.
- *Up to 128-bit Data Width* – Supports data paths up to 128 bits. The Avalon interface supports data widths that are not an even power of two.
- *Synchronous Operation* – Provides an interface optimized for synchronous, on-chip peripherals. Synchronous operation simplifies the timing behaviour of the Avalon interface, and facilitates integration with high-speed peripherals.
- *Dynamic Bus Sizing* – Handles the details of transferring data between peripherals with different data widths. Avalon peripherals with differing data widths can interface easily with no special design considerations.
- *Simplicity* – Provides an easy-to-understand interface protocol with a short learning curve.
- *Low resource utilization* – Provides an interface architecture that conserves on-chip logic resources.
- *High performance* – Provides performance up to one-transfer-per-clock.

III. AVALON INTERRUPT INTERFACE

Avalon Interrupt interfaces allow slave components to signal events to master components. For example, a DMA controller can interrupt a processor when it has completed a DMA transfer [1]. It has two categories:

3.1 Interrupt Sender

An interrupt sender drives a single interrupt signal to an interrupt receiver. The timing of the irq signal must be synchronous to the rising edge of its associated clock, but has no relationship to any transfer on any other interface. irq must be asserted until the interrupt has been acknowledged on the associated Avalon-MM (Memory Mapped) slave interface. The interrupt receiver typically determines how to respond to the event by reading an interrupt status register from an Avalon-MM slave interface. The mechanism used to acknowledge an interrupt is component specific [1].

3.2 Interrupt Receiver

An interrupt receiver interface receives interrupts from interrupt sender interfaces. Components with an Avalon-MM master interface can include an interrupt receiver to detect interrupts asserted by slave components with interrupt sender interfaces. The interrupt receiver accepts interrupt requests from each interrupt sender as a separate bit [1]. An interrupt receiver interface receives interrupts from interrupt sender interfaces.

IV. PROPOSED DIAGRAM OF AVALON INTERRUPT INTERFACE

4.1 Individual Request

The interrupt receiver expects to see each interrupt request from each interrupt sender as a separate bit and is responsible for determining the relative priority of the interrupts. The Avalon Interrupt Interface with individual request is shown in Figure 1.

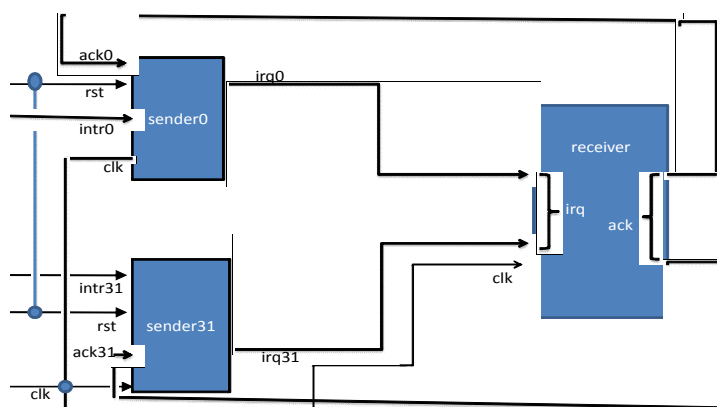


Figure1. Individual request diagram

4.2 Top Level Block Diagram

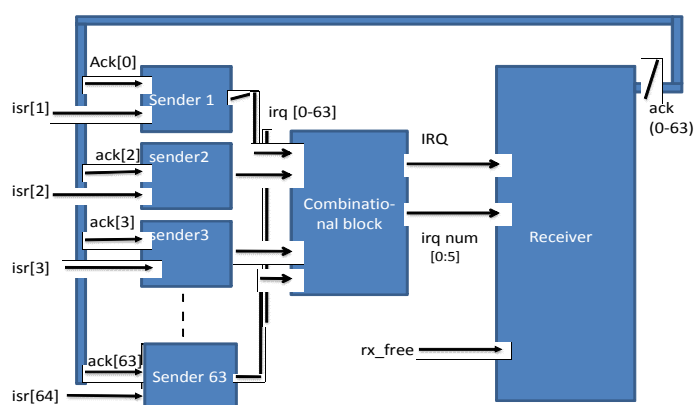


Figure2. Top level Block diagram

The block diagram of Avalon Interrupt Interface is shown in Figure 2. There can be maximum 64 senders.

4.3 Sender Block and Receiver Block

The sender and receiver block of Avalon Interrupt Interface is shown in Figure 3 and Figure 4. The input and output pin description of sender and receiver block has been explained in Table 1.

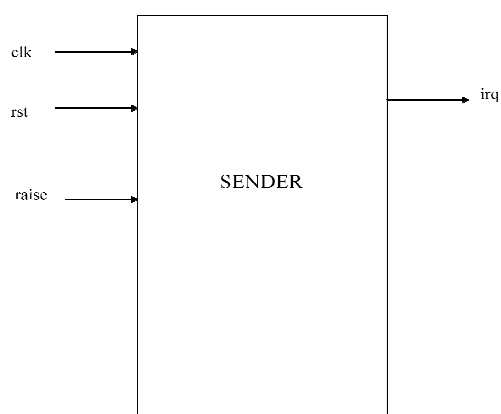


Figure 3. Sender block

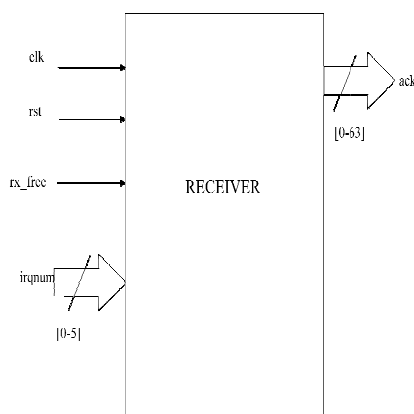


Figure 4. Receiver block

Table 1. Pin Description

Input Pin Description of Sender	
INPUT	Description
clk	Clock of 50 MHz
rst	Reset the program
raise	It is a single pulse signal
ack	Acknowledged signal from the Receiver
Output Pin Description of Sender	
OUTPUT	Description
irq	Interrupt request
Input Pin Description of Receiver	
INPUT	Description
clk	Clock of 50 MHz
rst	Reset the program
rx_en	Enable signal for the Receiver
Irq[63:0]	Interrupt request
Output Pin Description of Receiver	
OUTPUT	Description
Ack[63:0]	Acknowledged signal

4.3 State Diagram of Sender and Receiver Block

The state diagram of Sender and Receiver block is shown in Figure 5 and Figure 6.

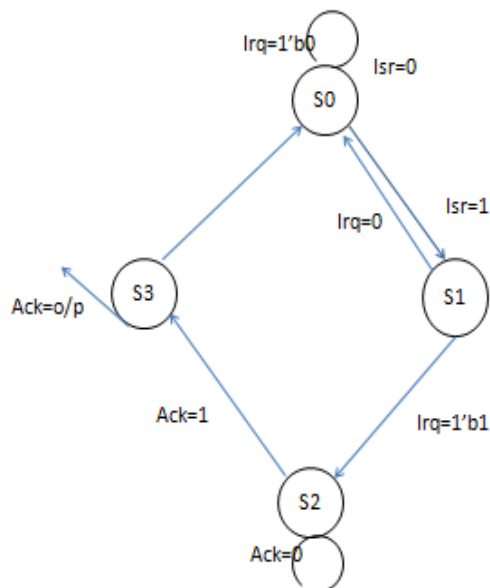


Figure 5.State diagram of Sender

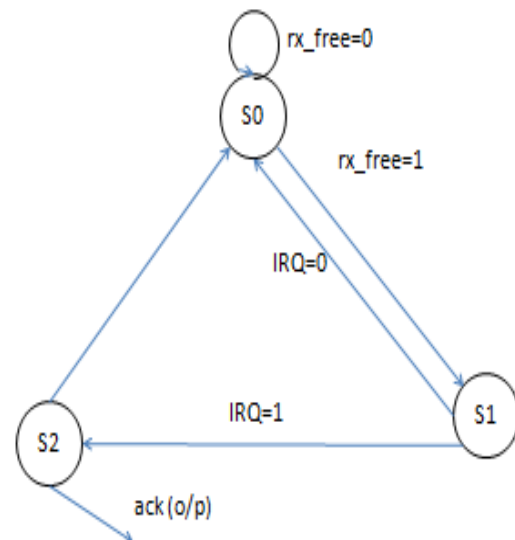
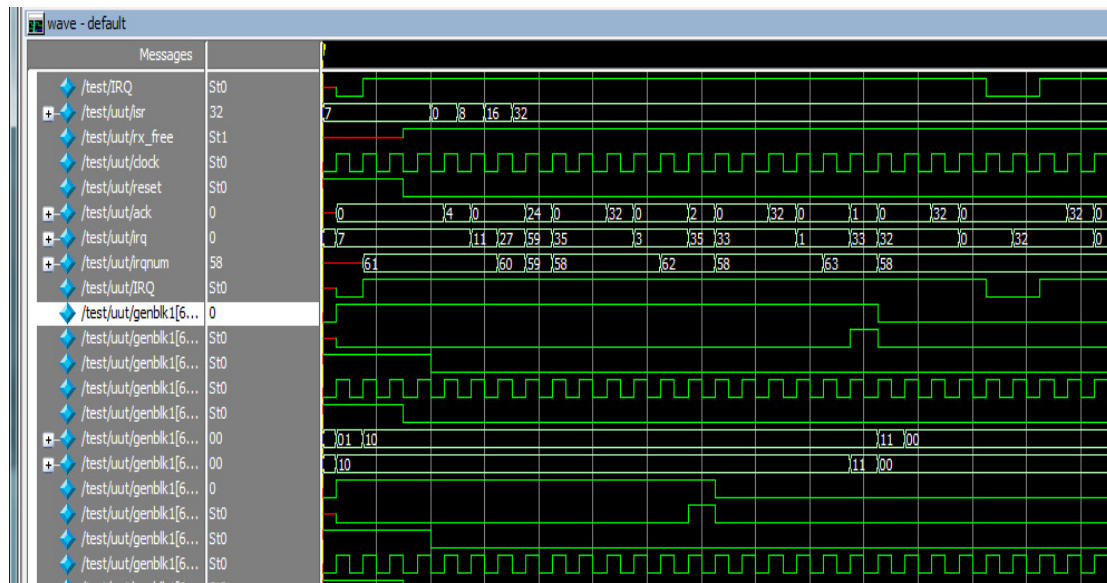


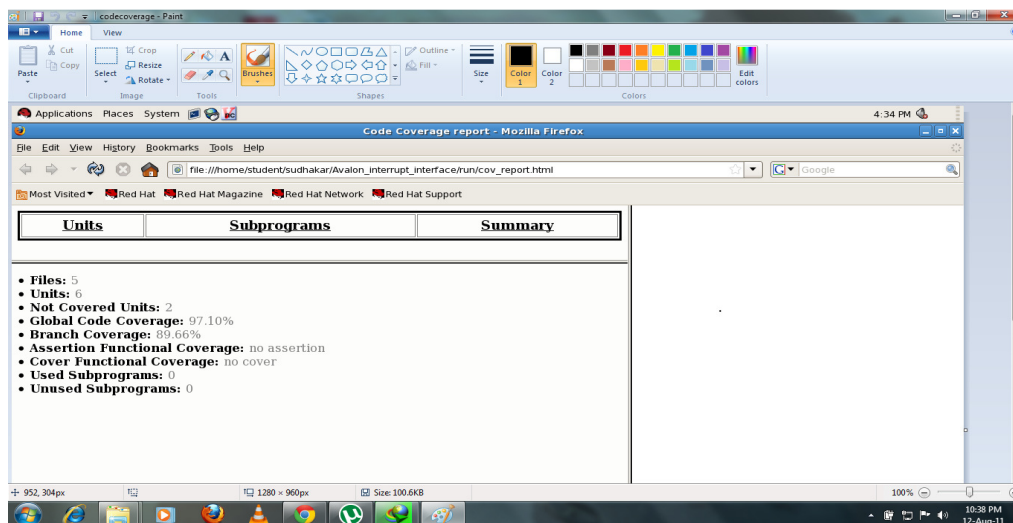
Figure 6.State diagram of Receiver

V. DUT WAVEFORM OF AVALON INTERRUPT INTERFACE

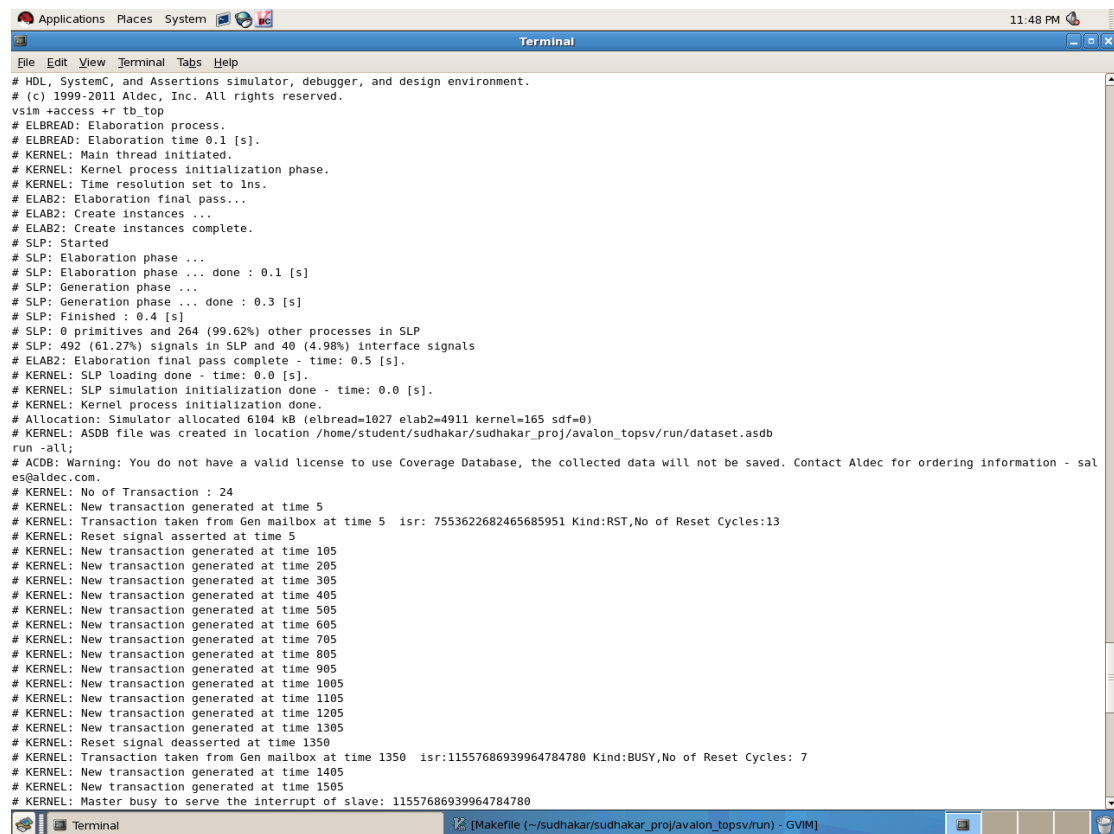
To measure the coverage analysis of design, the code was compiled and then simulated by using Questasim simulator to get the output which is shown in Figure 7.



VI. AVALON DUT CODE COVERAGE



VII. SIMULATION RESULT OF VERIFICATION OF AVALON INTERRUPT INTERFACE

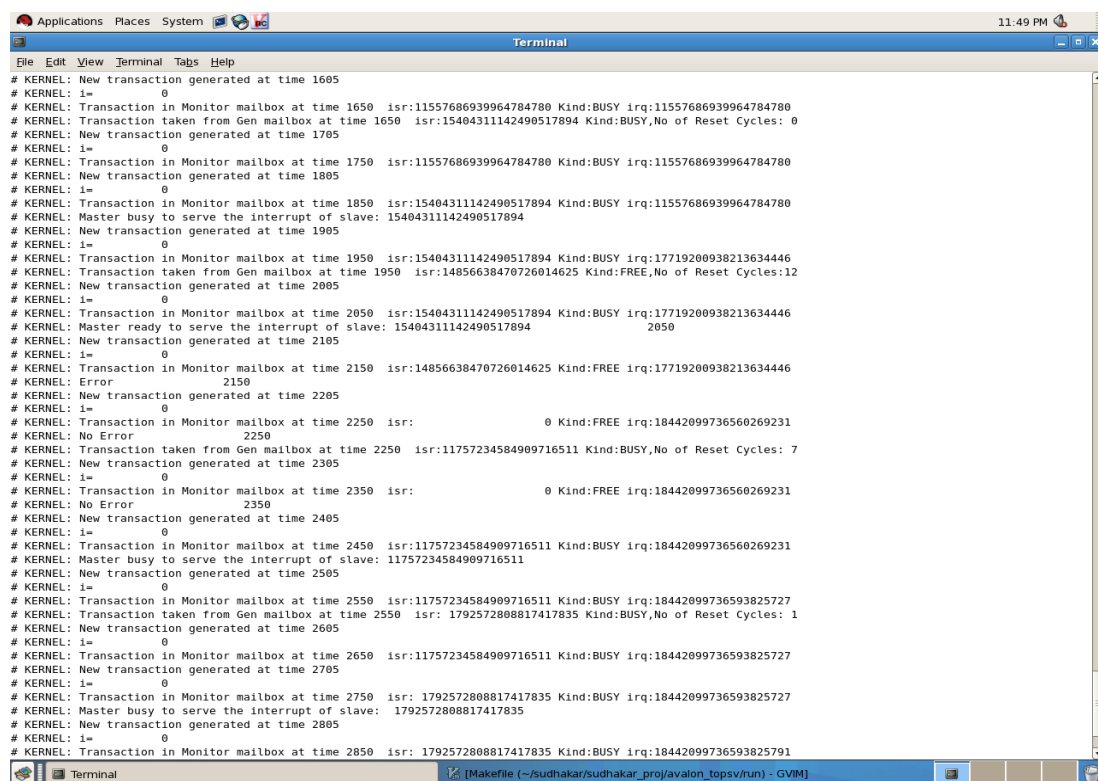


```

# HDL, SystemC, and Assertions simulator, debugger, and design environment.
# (c) 1999-2011 Aldec, Inc. All rights reserved.
vsim +access +r tb_top
# ELBREAD: Elaboration process.
# ELBREAD: Elaboration time 0.1 [s].
# KERNEL: Main thread initiated.
# KERNEL: Kernel process initialization phase.
# KERNEL: Time resolution set to 1ns.
# ELAB2: Elaboration final pass...
# ELAB2: Create instances ...
# ELAB2: Create instances complete.
# SLP: Started
# SLP: Elaboration phase ...
# SLP: Elaboration phase ... done : 0.1 [s]
# SLP: Generation phase ...
# SLP: Generation phase ... done : 0.3 [s]
# SLP: Finished : 0.4 [s]
# SLP: 0 primitives and 264 (99.62%) other processes in SLP
# SLP: 492 (61.27%) signals in SLP and 40 (4.98%) interface signals
# ELAB2: Elaboration final pass complete - time: 0.5 [s].
# KERNEL: SLP loading done - time: 0.0 [s].
# KERNEL: SLP simulation initialization done - time: 0.0 [s].
# KERNEL: Kernel process initialization done.
# Allocation: Simulator allocated 6104 kB (elbread=1027 elab2=4911 kernel=165 sdf=0)
# KERNEL: ASDB file was created in location /home/student/sudhakar/sudhakar_proj/avalon_topsv/run/dataset.asdb
run -all;
# ACDB: Warning: You do not have a valid license to use Coverage Database, the collected data will not be saved. Contact Aldec for ordering information - sal
es@aldec.com.
# KERNEL: No of Transaction : 24
# KERNEL: New transaction generated at time 5
# KERNEL: Transaction taken from Gen mailbox at time 5 isr: 7553622682465685951 Kind:RST,No of Reset Cycles:13
# KERNEL: Reset signal asserted at time 5
# KERNEL: New transaction generated at time 105
# KERNEL: New transaction generated at time 205
# KERNEL: New transaction generated at time 305
# KERNEL: New transaction generated at time 405
# KERNEL: New transaction generated at time 505
# KERNEL: New transaction generated at time 605
# KERNEL: New transaction generated at time 705
# KERNEL: New transaction generated at time 805
# KERNEL: New transaction generated at time 905
# KERNEL: New transaction generated at time 1005
# KERNEL: New transaction generated at time 1105
# KERNEL: New transaction generated at time 1205
# KERNEL: New transaction generated at time 1305
# KERNEL: Reset signal deasserted at time 1350
# KERNEL: Transaction taken from Gen mailbox at time 1350 isr:11557686939964784780 Kind:BUSY,No of Reset Cycles: 7
# KERNEL: New transaction generated at time 1405
# KERNEL: New transaction generated at time 1505
# KERNEL: Master busy to serve the interrupt of slave: 11557686939964784780

```

Figure 9. Simulated results of Verification of Avalon interrupt Interface



```

# KERNEL: New transaction generated at time 1605
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 1650 isr:11557686939964784780 Kind:BUSY irq:11557686939964784780
# KERNEL: Transaction taken from Gen mailbox at time 1650 isr:15404311142490517894 Kind:BUSY,No of Reset Cycles: 0
# KERNEL: New transaction generated at time 1705
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 1750 isr:11557686939964784780 Kind:BUSY irq:11557686939964784780
# KERNEL: New transaction generated at time 1805
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 1850 isr:15404311142490517894 Kind:BUSY irq:11557686939964784780
# KERNEL: Master busy to serve the interrupt of slave: 15404311142490517894
# KERNEL: New transaction generated at time 1905
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 1950 isr:15404311142490517894 Kind:BUSY irq:17719200938213634446
# KERNEL: Transaction taken from Gen mailbox at time 1950 isr:14856638470726014625 Kind:FREE,No of Reset Cycles:12
# KERNEL: New transaction generated at time 2005
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2050 isr:15404311142490517894 Kind:BUSY irq:17719200938213634446
# KERNEL: Master ready to serve the interrupt of slave: 15404311142490517894 2050
# KERNEL: New transaction generated at time 2105
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2150 isr:14856638470726014625 Kind:FREE irq:17719200938213634446
# KERNEL: Error 2150
# KERNEL: New transaction generated at time 2205
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2250 isr: 0 Kind:FREE irq:18442099736560269231
# KERNEL: No Error 2250
# KERNEL: Transaction taken from Gen mailbox at time 2250 isr:11757234584909716511 Kind:BUSY,No of Reset Cycles: 7
# KERNEL: New transaction generated at time 2305
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2350 isr: 0 Kind:FREE irq:18442099736560269231
# KERNEL: No Error 2350
# KERNEL: New transaction generated at time 2405
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2450 isr:11757234584909716511 Kind:BUSY irq:18442099736560269231
# KERNEL: Master busy to serve the interrupt of slave: 11757234584909716511
# KERNEL: New transaction generated at time 2505
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2550 isr:11757234584909716511 Kind:BUSY irq:18442099736593825727
# KERNEL: Transaction taken from Gen mailbox at time 2550 isr: 1792572808817417835 Kind:BUSY,No of Reset Cycles: 1
# KERNEL: New transaction generated at time 2605
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2650 isr:11757234584909716511 Kind:BUSY irq:18442099736593825727
# KERNEL: New transaction generated at time 2705
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2750 isr: 1792572808817417835 Kind:BUSY irq:18442099736593825727
# KERNEL: Master busy to serve the interrupt of slave: 1792572808817417835
# KERNEL: New transaction generated at time 2805
# KERNEL: i= 0
# KERNEL: Transaction in Monitor mailbox at time 2850 isr: 1792572808817417835 Kind:BUSY irq:18442099736593825791

```

Figure 10. Simulated results of Verification of Avalon interrupt Interface

VIII. COVERAGE ANALYSIS OF VERIFICATION OF AVALON INTERRUPT INTERFACE

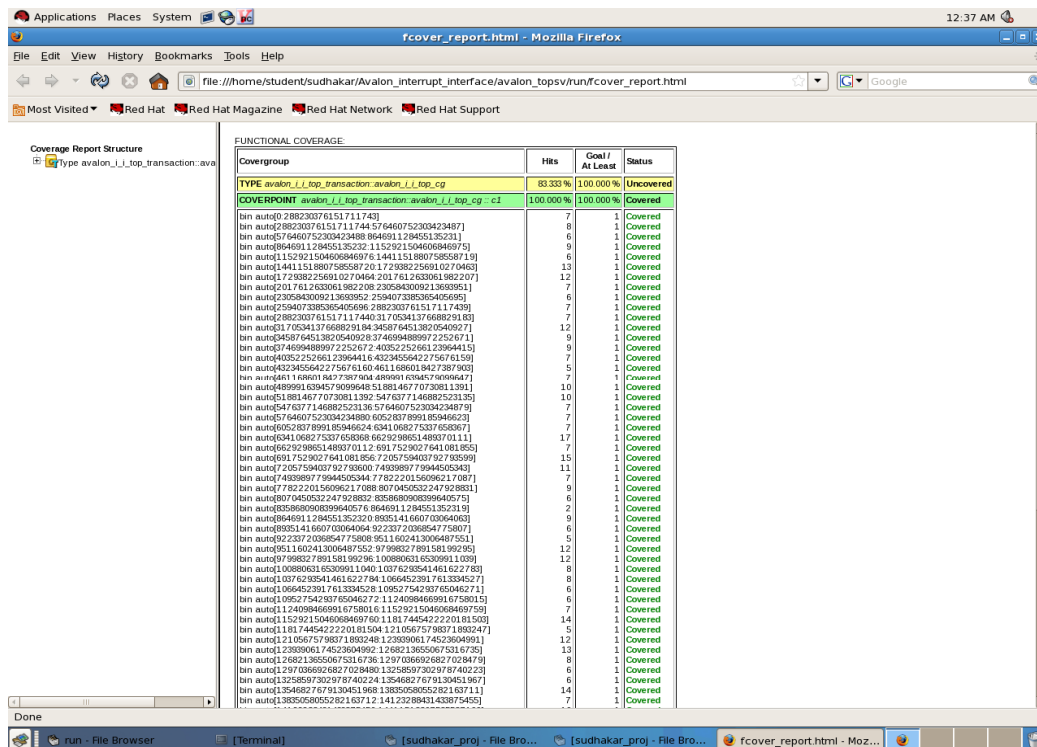


Figure 11. Coverage results of Verification of Avalon interrupt Interface

The Coverage Report gives the details of the functional coverage when complete analysis was done for the DUT and generated coverage report is shown in Figure 6. It is found that the coverage is 100%.

IX. CONCLUSION

The VMM Standard Library provides the foundation base classes for building advanced testbenches, while VMM Application provide higher level function for improved productivity. In this paper Avalon Interrupt Interface IP and VMM complaint VIP is developed. The IP was subjected to various analyses. It was observed that the simulated result of the Avalon Interrupt Interface IP was perfect and code coverage report was also satisfactory. The IP was verified for code coverage using QuestaSim. During the process of code coverage it was found that though the IP gave the expected output, still its overall code coverage is not 100% or approximately nearer to it, indicating that all the paths of the design has not been used to their full strength. So it is clear that higher order exhaustive verification has to be taken to reduce the errors in the design and increase the efficiency of Avalon Interrupt Interface IP. Thus, IP was subjected to VIP developed using VMM environment and was verified exhaustively. The result generated for VIP was found satisfactory and hence Avalon Interrupt interfaces can be used to allow slave components to signal events to master components.

REFERENCES

- [1] "Avalon Interface Specifications", Altera Corporation, May 2011.
- [2] "Avalon Interface Specification", Altera Corporation, April 2005.
- [3] Akhilesh Kumar and Mahesh Kumar Jha, "Coverage Analysis in Verification of Total Zero Decoder of H.264 CAVLD" IJAET Journal, Vol. 1, Issue 5, pp. 197-203, Nov 2011.
- [4] Akhilesh Kumar and Chandan Kumar, "Functional Coverage Analysis of OVM Based Verification of H.264 CAVLD Slice Header Decoder" IJAET Journal, Vol. 1, Issue 5, pp. 109-117, Nov 2011.
- [5] Chris Spear, SystemVerilog for Verification, New York : Springer, 2006.
- [6] Bergeron, Janick. Writing testbenches: functional verification of HDL models. s.l.: Springer, 2003.
- [7] Mihaela E.Radhu and Shannon M.Sexton, "Integrating Extensive Functional Verification into digital design Education," IEEE Trans. Educ., vol. 51, no. 3, pp. 385–393, Aug.2008

- [8] VMM Golden Reference Guide ,Vers. 1.2, DOULOS, January 2010, Available: www.vmmcentral.org
- [9] VMM for SystemVerilog Companion Guide, Vers. 1.1, DOULOS, December 2007, Available: <http://www.doulos.com>
- [10] Samir Palnitkar, "Verilog® HDL: A Guide to Digital Design and Synthesis", Second Edition
- [11] <http://www.altera.com/support/software/system/sopc/sof-sopc-avalon-interfaces.html>
- [12] <http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.dui0395b/CJHHCEAE.html>
- [13] www.ict.kth.se/courses/SystemVerilogFunctionalCoverage.pdf
- [14] <http://www.testbench.co.in>

Authors

Mahesh Kumar Jha received B.Tech. Degree from Biju Patnaik University of Technology, Orissa, India in 2007. Currently he is pursuing M. Tech in Department of Electronics and Communication Engineering at N.I.T. Jamshedpur, Jharkhand, India. His interested field of research is VLSI Design.



Richa Sinha received B. E. Degree from RajaRamBapu Institute of Technology Shivaji University, Kolhapur, Maharashtra, India in 2007. Currently she is pursuing M. Tech project work under the guidance of Prof. Akhilesh Kumar in the Department of Electronics & Communication Engg, N. I. T., Jamshedpur. Her interest of field is ASIC Design & Verification.



Akhilesh Kumar received B.Tech degree from Bhagalpur University, Bihar, India in 1986 and M.Tech degree from Ranchi, Bihar, India in 1993. He has been working in teaching and research profession since 1989. He is now working as H.O.D. in Department of Electronics and Communication Engineering at N.I.T. Jamshedpur, Jharkhand, India. His interest of field of research is analog and digital circuit design in VLSI.



CONCATENATION OF BCH CODE WITH SPACE TIME CODE A LOW SNR APPROACH FOR COMMUNICATION OVER POWER LINES FOR SUBSTATION AUTOMATION

Rajeshwari Itagi¹, Vittal K. P.², U. Sripathi³

¹Department of E & C, KLE Institute of Technology, Hubli, Karnataka, India.

²Department of E&E., NITK, Surthkal, Karnataka, India.

³Department of E&C., NITK, Surthkal, Karnataka, India.

ABSTRACT

A study of power line as communication medium for narrowband application is performed in the perspective of space time coding. The spatial diversity available in the naturally decoupled phases of the power line can be exploited for using power line as communication medium for narrowband application just as incase of fading channel of a wireless communication. In this paper, a study is performed to test the transmission of digital data using simulated power line. Time and frequency dependent transfer function of the power line is varied instantaneously during data transmission, which is a characteristic of the power line. Impulsive disturbance on the power line is modeled using Middleton class-A noise. BCH code is used as outer code to reduce the SNR requirement. Concatenation of BCH code with space time code is found to reduce the SNR requirement to have the carrier power within the regulatory limits.

KEYWORDS: Space time coding, Multipath channel, Power line channel, Impulse noise, BCH coding

I. INTRODUCTION

Power Line Communication (PLC) refers to communication of information with power line as the medium for communication. Digital communication has replaced the former analog PLC. Digital data corresponding to information is encoded first with error control coding and then modulated using digital modulation schemes. The modulated data being continuous in time is coupled to the high voltage power line by means of a coupling transformer in series with a capacitor. Similar capacitor transformer pair at the receiver point collects the transmitted data and information is recovered that is useful to operate the receiver device.

The high frequency carrier signal carrying the digital data information when coupled to the power line carrying power with low frequency (50-60 Hz), the line tends to act as antenna and radiates electromagnetic (EM) waves. This feature of EMI (Electro Magnetic Interference) of the line is undesirable, as the EM radiations may trigger or distort the unintended receivers operating in the same frequency range. Therefore regulatory standards put the limit on the carrier power to be injected on the line. CENELEC EN 50065-1 (committee European de Normalization Electro technique (CENELEC), regulation EN50065-1) allows 90-148 kHz and FCC (Federal Communication Commission) allows up to 500 kHz, with a permissible carrier limit of 134 dB μ V for industry automation [4].

Different approaches are attempted to reduce the carrier power used for digital communication on power line, in order to reduce the EMI. In this paper, a scheme for modem (modulator demodulator) design is derived to achieve the goal of carrier signal power reduction, for data communication over power lines, for the purpose of substation automation is discussed.

Power line channel for data transmission follows two classifications as

(i). Wideband or Broadband over Power Line Communication (BPLC), to transmit and receive internet data which is further classified as:

- (a) Low voltage (<1 kV) high speed BPLC (>2Mbps) for internet application
- (b) High voltage (>36kV) high speed BPLC (>2Mbps) for internet application

(ii) Narrowband Power Line Communication (NPLC) intended for supervisory or control data transmission on power line which is further classified as:

- (a) Low voltage low speed NPLC for home automaton (<100kbps)
- (b) Medium voltage (1kV – 30 kV) low speed NPLC for industry automation

A substation is said to be automated when it employs intelligent electrical/electronic devices (IED's) for supervisory and control actions. Substation automation can be achieved by communication with IED's using power line. Thus substation automation comes under the class of narrowband PLC. Carrier frequency and carrier power used for digital modulation of the data should be within the regulatory limits specified for NPLC.

In [1], the author mentions the merits of Medium voltage PLC for automation such as emergency and maintenance control, security systems, power network (power grid) management optimization and monitoring systems such as remote metering, power quality measurement and fault survey.

In [1, 2, 3 and 4], authors have explained that the power line channel can be modeled as a multipath channel, as in case of wireless or mobile channel. A multipath channel is also called as fading channel, due to characteristic of fading phenomenon of multipath channel. Fading is the property of instantaneous variation of signal power in a multipath channel.

In [2], the author suggests the idea of exploiting naturally decoupled phases of the power line and power line being referred to as fading channel, so that space time coding can be applied to power line, to use power line as communication medium for NBPL. The simplest of space time code requires two transmitters and one receiver that can be achieved by coupling the same signal to two phases of the line and collecting from one phase.

The use of power line as medium of communication is attractive as the channel is ready to use and no new medium is required to be created and hence cost effective. The impulsive noise present on the line makes the PLC as not so reliable and this requires the use of efficient schemes of modulation and channel coding.

Alamouti's two transmit antenna one receive antenna space-time code [7] has been adopted in various global standards and is used to estimate and detect the modulated data in a fading channel. Space time coding is applied to PLC in this paper to predict the possibly highly corrupted information on the line and to reduce carrier power. Alamouti's code requires the knowledge of channel state or channel transfer function, which is carried out by sending dummy pilot symbols.

Section 2 of this paper is arranged to provide the nature of the transfer function of the power line channel and the associated impulse noise and the details to model them. Space time coding scheme is explained in section 3. Results are discussed in section 4. Conclusions are given in section 5.

II. POWER LINE CHANNEL AND IMPULSE NOISE

PLC is classified as broadband PLC (BPLC) for high speed internet data communication and as low speed narrowband PLC (NPLC) for industry or home automation data communication. Modeling the power line channel is carried out by two main types- multipath model and by using ABCD parameters by different authors. Modeling of power line in this paper is done by multipath model [5].

The power line can be modeled as multipath channel [1, 5]. Any transmission line will have reflections on the line, if impedance mismatches occur. Impedance mismatch is obviously present on line as few tapping points on the power line always switch 'on' or 'off'. The 'on' or 'off' of switching the load points cause the line impedance impairment and causing reflections or multipaths. This makes the channel to be modeled as multipath channel. The magnitude response of channel thus varies with respect to time as well as with frequency making it to be referred to be as time and frequency variant channel [1, 3 and 5].

$$H(f) = \sum_{i=1}^N g_i \cdot A(f, d_i) \cdot e^{-j2\pi f\tau_i} \quad (1)$$

$$H(f) = \sum_{i=1}^N g_i \cdot e^{-(a_0 + a_1 \cdot f^k) \cdot d_i} \cdot e^{-j2\pi f \tau_i} \quad (2)$$

Equation (1) gives the transfer function of the multipath channel and equivalent to (1), (2) gives transfer function of power line channel [5]. g_i in (1) and (2) represents the gain of each path at i^{th} instant, exponent of second term in (2) represents the attenuation of $H(f)$ with respect to frequency and the exponent of third term in (2) represents the delay (phase changes) in the received multipaths at i^{th} instant. N represents the number of multipaths in wireless channel given by (1) and number of disturbance points in a power line channel, given by (2). Thus (2) represents the time variant and frequency variant transfer function of power line channel. d_i are lengths of disturbance points along the power line and dielectric constant of the insulating material is represented by ϵ_r . The pdf (probability density function) of Middleton's Class-A Noise, is given by (4).

$$f(x) = \sum_{m=0}^{\infty} \frac{A^m}{m!} e^{(-A)} \left[\frac{1}{\sqrt{2\pi\sigma_m^2}} \right] \cdot e^{\left(\frac{-x^2}{2\sigma_m^2} \right)} \quad (4)$$

Equation (4) suggests that Middleton Class A noise model [6] that refers to cumulative sum of Gaussian distributions where σ_m^2 is noise variance, with $\sigma_m^2 = \sigma^2 (m/A+T) / (1+T)$, index $T = \sigma_G^2 / \sigma_{GI}^2$ is the GIR (Gaussian-to-Impulsive noise power Ratio) with Gaussian noise power σ_G^2 and impulsive noise power σ_{GI}^2 . And $\sigma^2 = \sigma_G^2 + \sigma_{GI}^2$ is the total noise power. The noise x followed by (4) always includes the background Gaussian noise with power σ_G^2 .

In [4], several noise sources that can be found on low- or medium-voltage power grids are explained, such as, for example (i) colored thermal noise, (ii) periodic asynchronous impulse noise related to switching operations of power supplies, (iii) periodic synchronous impulse noise mainly caused by switching actions of rectifier diodes. Reference to different literature, suggests that analysis of power line with impulse noise justifies the check for power line communication performance.

III. SPACE TIME CODING

A multipath channel is also called as fading channel, due to characteristic of fading phenomenon of multipath channel. Fading is the property of instantaneous variation of signal power in a multipath channel.

3.1 Transmit and Receive scheme for space time coding

In [7], inventions of S.M. Alamouti, the so-called Alamouti space-time block codes, filed in 1997 and patented jointly with Vahid Tarokh are explained. Alamouti's code is a 2 transmit antenna space-time code and has been adopted in various global standards.

Figure 1 gives the schematic of space time coding using 2X1 system (2 transmit antennas, 1 receive antenna). Table I gives the transmit scheme for communication system shown in Fig.1

s_0, s_1 are symbols transmitted in two time slots as per Table 1. r is received signal $[r_0 \ r_1]$, H_0 and H_1 are transfer functions of channel 1 and channel 2 respectively, also called as channel states. s_{0e} and s_{1e} are estimates of symbols s_0 and s_1 at time slots t_0 and t_1 at the receiver which are determined by (5) and (6).

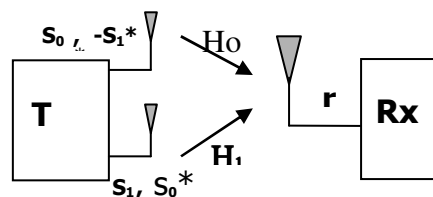


Figure 1. Two transmit one receive space time coding

Table I. Transmit scheme for space time coding

	Time instant t_0	Time instant t_1
Transmitter point 1	symbol 1	-Conj(symbol2)
Transmitter point 2	symbol 2	Conj(symbol1)

$$s_{oe} = \text{conj}(H_0) \times r_0 + H_1 \times \text{conj}(r_1) \quad (5)$$

$$s_{ie} = \text{conj}(H_1) \times r_0 - H_0 \times \text{conj}(r_1) \quad (6)$$

Recovery of symbols s_o and s_i are further found by minimum distance of s_{oe} and s_{ie} with symbols s_o and s_i . The detailed mathematical support to derive receiving scheme is worked in [7]. It is seen that recovery at the receiver requires the knowledge of channel states. Signal estimation and detection at the receiver requires the knowledge of the channel states (Channel State Information). Channel estimation is performed by adding known dummy pilot symbol.

3.2 Power Line as multipath fading channel

The two transmitting antennas in the Alamouti's space time coding scheme can be realized in power line channel, by injecting modulated signal to two different phase lines of power line with respect to earth line by means of capacitive coupled coupling transformers and then by tapping received signal between two phase lines. Equivalent to two fading paths in a wireless channel, the two different paths chosen for signal transmission in power line are statistically uncorrelated and thus make it possible to use the space time scheme.

IV. RESULTS AND DISCUSSION

As per the theory of space time coding applied to fading channel, when signal deteriorates in one path, but is not severely disturbed by any one of other paths, then there should be proper recovery. The same is verified from the results obtained.

The data symbols from information source are first protected by error control coding. BCH code (127, 22), [8] is used to channel code the data. Symbols are then modulated by phase shift keying (PSK) and a carrier frequency within the permissible frequency range used for narrowband PLC. The modulated data then is sent using space time coding as per Table I, and recovered as explained in section 3.

Fig. 2 and Fig. 4 show the plot of magnitude of transfer function in dB (attenuation in dB) of the simulated power line for two different paths. The two paths here mean that one path corresponding to signal propagating via phase 1 of line and path 2 corresponds to signal propagating via phase2 of line.

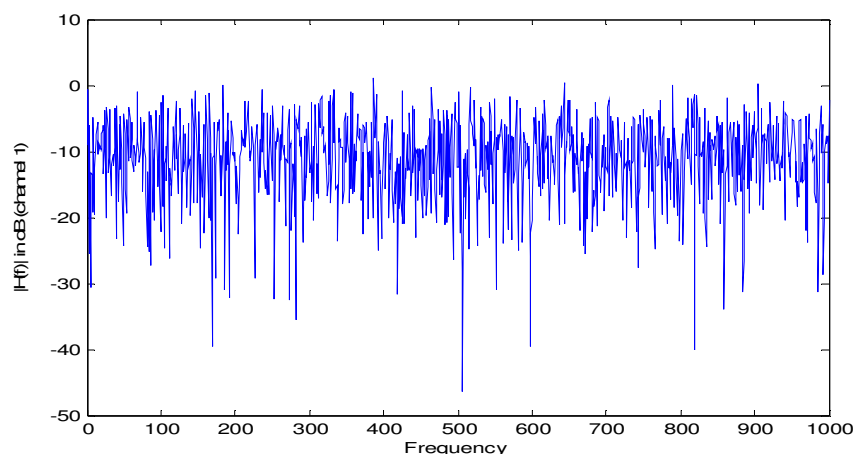
**Figure 2.** Magnitude of transfer function in path 1.

Table II. Range of $|H(f)|$ in dB with type of attenuation assumed

Range of $ H(f) $ in dB	Attenuation
0 to -5	very low
-5 to -15	low
-10 to -45	medium
-30 to -80	high
-40 to -130	very high

Matlab 7.9 is used to simulate the proposed scheme. Transfer functions of two paths in the power line, are realized using (2), using different attenuation. Table II gives the range of values of magnitude of transfer functions referred to as low, very low, medium, high and very high attenuation, assumed in this paper.

Attenuation at the operating frequency (carrier frequency) is made to change by varying a_0 , a_1 and k , d_i and g_i . Frequency range of 90 to 500 kHz is considered for realizing the channels. The carrier frequency of 200 kHz is used to modulate the data in this paper.

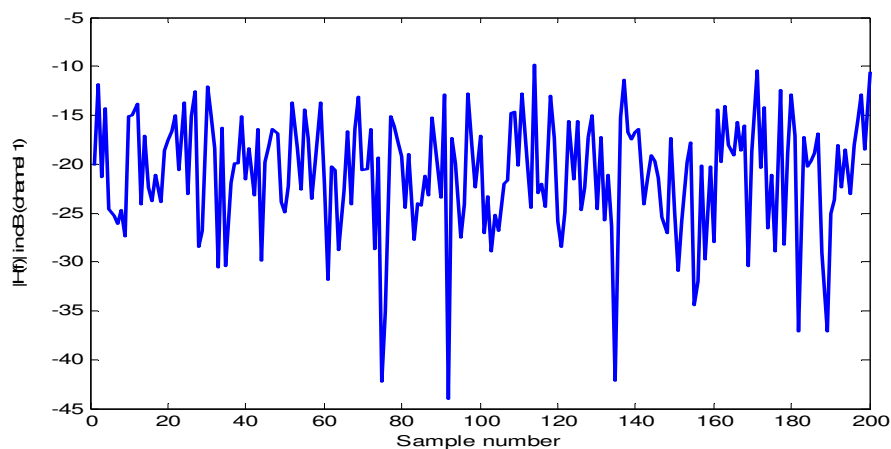


Figure 3. Magnitude of transfer function for path 1 for 200 bits.

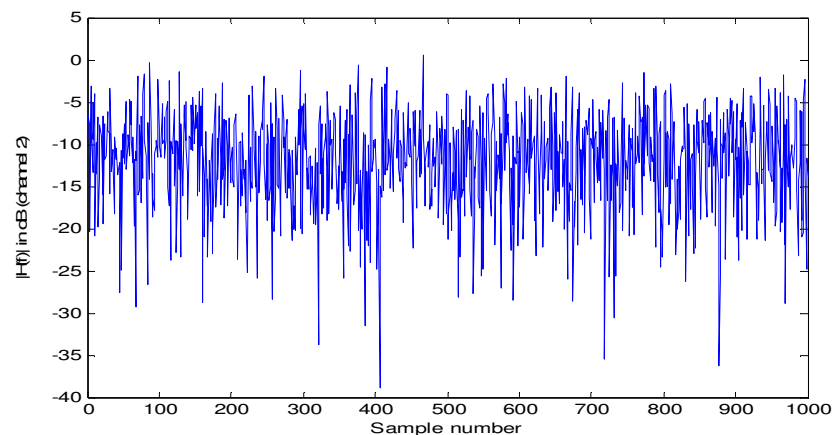


Figure 4. Magnitude of transfer function in path 2.

Time samples of impulse noise simulated using (4) are added to the signal transmitted with channel. In one path going high while the other path being with low and medium attenuation and reversing the situation next time. Data size of the order of 10^4 to 10^5 is used to test the results. Values of A and T are set to 0.1. Numbers of multipaths vary from 2 to 10, with distance between transmitter and receiver being 50 to 200m.

When the digital data is modulated with PSK, with assumed carrier of 200 kHz, the attenuation offered to every bit is shown in Figure. 2. In Figure. 3, attenuation for only 200 bits is shown for clarity, which can be shown for other bits similarly. Attenuations shown by Figure. 2 and Figure. 4,

can now be regarded as two different uncorrelated paths of power line channel. The bits recovered from space time decoding and BCH decoding, are compared with transmitted bits and a graph of probability of symbol (bit) error versus SNR (Signal to Noise Ratio) is plotted. The recovery of symbols is studied for different amounts of attenuation on two paths and Table III was prepared, with reference to Table II.

Table III Consolidated result of the scheme

Data size	A	T	d	No. of paths	Transfer Function in dB		Nature of channel transfer function	Value of SNR in dB for $P_e < 10^{-4}$
					$ H_1(f) $ indB	$ H_2(f) $ indB		
10^4	0.1	0.1	1- 2.5 m	4	Very High	Low	Constant for all bits	13
10^4	0.1	0.1	1- 2.5 m	4	Very High	Medium	Constant for all bits	19
10^5	0.1	0.1	1- 2.5 m	4	Medium	Low	Varying for every symbol	9
10^5	0.1	0.1	1- 2.5 m	4	Medium	High	„	13
10^5	0.1	0.1	1- 2.5 m	4	Medium	Medium	„	11
10^4	0.1	0.1	1- 2.5 m	4	Medium	High	„	13
10^5	0.1	0.1	1-200 m	6	Very High	Medium	„	17
10^5	0.1	0.1	1-5m	6	Medium	Medium	„	15
10^5	0.1	0.01	1-5 m	6	Medium	Medium	„	11
10^5	0.01	0.01	1-5 m	6	Medium	Medium	„	9
10^5	0.1	0.1	1-100 m	14	High	Medium	„	13
10^5	0.01	0.01	1-100 m	14	High	Medium	„	13

4.1 Channel state estimation

The reliability and performance of the proposed scheme depend on successful channel estimation.

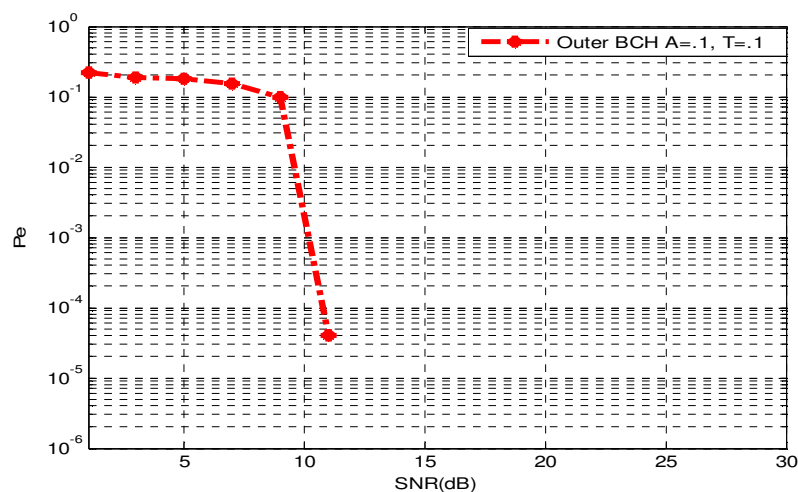


Figure 5. P_e vs. SNR for $A=0.1$, $T=0.1$, 10^5 bits

The knowledge of transfer functions H_0 and H_1 are required at the receiver for signal recovery as given by (5) and (6).

In most of the literature, CSI is assumed. i.e., values of H_0 and H_1 used to send data at transmitter are used as they are, at the receiver. CSI is not assumed as known in this paper. CSI is derived from dummy pilot symbols put along with data symbols.

V. CONCLUSIONS

Table III gives the complete work details and results obtained in this paper. Table III helps to verify the successful data transmission over power line using space time code and BCH code. Error performance of the order of 10^{-5} to 10^{-6} is observed within 20 dB of SNR. Carrier signal level computed at this SNR is found to be within specified limits for narrow band home/industry automation PLC application. In Figure 2, attenuation in path 1, with transfer function $|H_0(f)|$ is varying between 15 to 45 dB and attenuation in path 2, with transfer function $|H_1(f)|$ is varying between 0 to 40 dB, in figure 4. Transfer function for path 1 is shown for only 200 symbols (bits), to view the path attenuation with clarity in figure 3. Attenuation in path 1 and in path 2 can be regarded as medium for industry/home automation PLC application.

Error due to BCH coding which is in outer layer to Alamouti's space time code (inner layer coding), falls rapidly, after few initial values. This is because; the inner code has brought the error in the limits of complete error correction capacity of outer BCH code.

Results show that the system can tolerate the high magnitude impulses and system works well in protecting data from the impulse noise. There are many a versions of modem design using space time coding as in [2], which uses ideal channel state information (CSI). In the work presented in this paper, ideal channel state information is not assumed, but is derived from pilot symbols placed within the data bits. When ideal channel is assumed in the simulation studies, it is found that P_e will be less than 10^{-6} within 3 dB (not shown by plot), for $A=0.1$ and $T=0.1$. This shows that there is scope for improvement in the proposed modem design, if an improved CSI scheme is adopted. To find CSI, many a schemes use a learning sequence, for a length of few bits (or symbols) which may create unnecessary delay or increase in overhead. The proposed scheme does not use a long learning sequence. Also there will be no synchronization problems in this scheme, as the symbol estimation is performed using space time decoding and detection by maximum Euclidian distance in the signal space.

If one of the channel is having low or medium attenuation, then high attenuation in the other channel is tolerated and this holds good alternatively. This verification goes well with Alamouti's theory [7]. i.e., as per the theory of space time coding applied to fading channel, when signal deteriorates in one path, but is not severely disturbed by any one of other paths, then there should be proper recovery.

Error performance of the order of 10^{-5} is observed within 20 dB of SNR. Carrier signal level computed at this SNR is found to be within specified limits ($<134\text{dB}\mu\text{V}$) for narrow band power line communication.

Input of different data sizes are verified to provide the consistent results as tabulated in Table III. Impulse noise parameters are changed to check the error correction. Path lengths in the channel model and the number of paths are changed to verify error correction. Thus an efficient modem design for narrowband PLC is derived.

REFERENCES

- [1]. A Cataliotti, D D Cara, G Tinè, "Model of Line to Shield Power Line Communication System on a Medium Voltage Network", 2010, IEEE Transactions
- [2]. A. Cataliotti, A. Daidone, G. Tinè, "Power line Communications in Medium Voltage system: Characterization of MV cables", IEEE Transactions on Power Delivery, vol. 23, October 2008.
- [3]. A. Cataliotti, A. Daidone, G. Tinè, "A Medium Voltage Cable model for Power Line Communication", IEEE Transactions on Power Delivery, vol. 24, pp. 129-135, January 2009.

- [4]. Luis F. Montoya, "Power Line Communications Performance Overview of the Physical Layer of Available protocols", IEEE Transactions.
- [5]. Halid Hrasnica, et. al., "Broad band Powerline Communications Networks, Network Design", John Wiley & Sons, 2004.
- [6]. Leslie A. Berry, "Understanding Middleton's Canonical Formula for Class-A Noise", IEEE Transactions on Electromagnetic compatibility, Vol. EMC-23, No.4, Nov 1981.
- [7]. S.M. Alamouti (October 1998). "A simple transmit diversity technique for wireless communications", IEEE Journal on Selected Areas in Communications Vol. 16, No. 8, Oct. 1998, pp 141–1458.
- [8]. Shu Lin, D J Costello, "Error Control Coding Fundamentals and Applications", Second Edition , Prentice Hall, 2003.

Authors

Rajeshwari L Itagi received her B.E. degree in Electronics and Communication from GIT, Belgaum, and Karnataka, India in 1985 and she is presently working as Assistant Professor in KLEIT, Hubli, Karnataka, India in the Department of E and C Engineering. She is presently pursuing Ph. D. research program at National Inst. of Technology, Karnataka, and Surthkal, India. Her current research focuses on modem design for narrowband power line communication to minimize EMI.



K P Vittal, received his PhD from Mangalore University in 1999. He is currently Professor and Head in the Department of Electrical and Electronics Engineering, National Institute of Technology Karnataka, Surthkal, India. His current research interest includes power system protection, power quality and design of embedded systems.



U. Sripathi received his Ph. D. from IISc, Bangalore in 2005. He is currently Associate Professor in the Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Surthkal, India. His current research includes error control coding and MIMO systems.



UTILIZATION OF EXTRUSION AS AN ADVANCED MANUFACTURING TECHNIQUE IN THE MANUFACTURE OF ELECTRIC CONTACTS

Virajit A. Gundale¹, Vidyadhar M. Dandge²

¹Prof. & Academic-Incharge, Sharad Institute of Technology, College of Engineering,
Yadrav, Kolhapur India.

²Asstt. Prof., Dr. J. J. Magdum College of Engineering Jaysingpur, Kolhapur, India

ABSTRACT

It is common that each engineering assembly product has A,B and C class components which together comprises assembly. A class component has high contribution in product function, cost and profitability of the engineering enterprise. A special technical research is must for this class of components. Product assembly is another crucial function where all failures are coming on the surface. Where all components are to be available simultaneously to complete assembly. This is important stage because if component supply is interrupted assembly function stops. Many times the components cannot be supplied due to selection and use of traditional manufacturing techniques used at the time of sample lot submission. These short cuts taken during development phase are becoming hurdles at the time of commercial or bulk production of assembly. Selection of most suitable and advance techniques for manufacturing of all components gives quality and quantity of all components and streamlines assembly functions. Keen approach towards manufacturing techniques eliminates quality and productivity problems which occur at the supply stage of the product. This also derives long term benefits such as customer satisfaction gain, product life cycle growth.

KEYWORDS: Manufacturing techniques, productivity, product life cycle

I. INTRODUCTION

This paper presents and discusses the results of the research work of application of advance manufacturing techniques used for long term elimination of quality and productivity problems of electrical contacts required for electro-mechanical power entry products. Due to highly competitive global market the time available for design, development and sample submission of product assembly is very short. Every manufacturing industry has to take each customer enquiry as an opportunity for business growth. After submission and approval of sample lot of new developing products the next immediate demand of the customer is bulk supplies of products. At this juncture the problem faced by manufacturing organization and they are unable to expedite the supplies as per customers schedules. This problem was taken for investigation after performing why-why analysis with related design and development team for this problem. We understood that the root cause for occurrence of this critical situation lies in the fact that is supplier has used traditional manufacturing techniques for supplying sample lots of brass electrical contacts. This product is socket assembly. The assembly has three varieties of components viz-Two types of plastic covers, three Brass contact of same specifications, and pair of mild steel fasteners. The BOM of the product indicates that Brass contact is a A class component and the supplies of these components are not made from Jamnagar based vendor is unable to supply contacts in time results in continuous stoppage of assembly results in business loss. This product is socket assembly of export market and brass electrical contact play key

role in the product assembly function. The samples are approved and at latter stage customer demands commercial supplies of the product but the product supplies cannot be rushed with required rates due to low productivity and quality of contacts at vendor end. Frequent occurrence of this situation ultimately resulted in the loss of customer satisfaction. This subject is very sensitive for the organization starving for business growth. This is so alarming that this product has further no future if this situation continues to remain the same. This has become major obstacle in export business growth of the organization. Plastic covers and fastener pair have no problems for quality and productivity. Hence to overcome this tricky business situation we focused on brass electrical contact manufacturing technology used by vendor and used hot extrusion technique followed drawing as advance metal forming techniques to give near and neat shape to T section of contact.

II. CONTACT SPECIFICATIONS

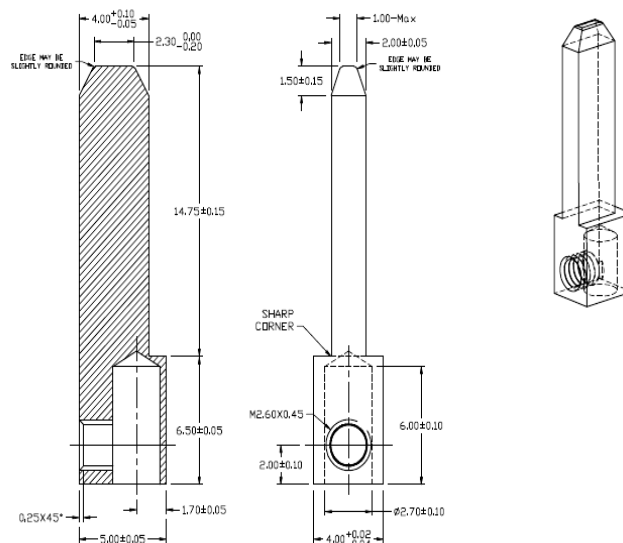


Figure. 1 Contact

This figure above shows sectional, front & isometric view of said contact. The basic material is 70-30 Brass.

2.1 Research Methodology The three phase research methodology is used to achieve desired results. Phases are as below.

Phase 1-Analyze existing Brass contact manufacturing technique and to find causes for low productivity and high rejections at various stages.

Phase 2 -To evaluate and compare probable alternative manufacturing technologies and select the most suitable one.

Phase 3-To select and establish most suitable advance manufacturing technique to come out with productivity and quality solutions for bottleneck contacts due to which the assembly work of the product is badly affected. This should solve organizational problem and rewarded new life to the product.

Existing technique is manufacturing raw casted contacts using green sand molding technology and machining all over. Technical evaluation of existing manufacturing technology this is done in two steps.

Step1-Evaluation of green sand molding and casting technology. Step2-Evaluation machining processes of the contacts to the finish dimensions.

The critical points observed as apart of casting process limitations are a) We must provide machining allowance of 2.50 mm all over the contact surfaces since we cannot get finish contacts by casting technique. Due to low sectional thickness (refer drawing) faster cooling of castings takes place results in hard scale formation b) For lower cross sections streamlining metal flow a big question mark. As a combined effect of these two heavy casting rejection occurs. This is as tabulated in table No 1

Table 1. Type of defect and rejection percentage

No	Type of defect	%
1	Pin holes and unfilled	12
2	Chilled surfaces	13
3	Scaling and blow holes	22

The total rejection is 47%. Existing sequence of operations of manufacturing of raw casted contacts is in total 14 operations required to get finish contact.

Table 2 indicates process and estimation of the cost incurred.

No	Operation	Time	Estimate Base.	Per contact. (INR)
1	Melting	On Wt. basis	Rs 30 /Kg. contact Wt. 10gm.	INR 0.30
2	Material cost.	Wt. basis	Rate INR 315/Kg	INR 3.15 (Raw Contact Wt. 10 gm.)
3	Design ,Dev. of pattern.	72 Hrs @250 INR/Hr=Rs 18000	One time cost.	INR 0.18 (Pattern life 100000 castings)
4	Mounting.	24 Hrs @INR 250.	6000 One time.	To amortize on 1000 molds ie 25000 castings =INR 0.24
5	Mold ing	100 /shift	INR 500/Shift	25 castings per mold. INR.0.13.
6	Pouring.	10 Min.	2 persons INR 500/shift	2=00 Divide by 25 No of patterns works INR 0.10
7	Knock out..	2 Min.	1 person	1=00/25 works to INR 0.04
8	Fettling and cleaning		On Kg basis	1=50 works to INR 0.15
9	Hardness testing.		On per piece basis	0=20 works to INR 0.02
10	Facing.	0.5 Min	INR 80/Hr.	INR 0.66
11	Straddle milling.	0.25 Min	INR 80/Hr	INR 0.33
12	Slot milling.	0.5 Min	INR80/Hr	INR 0.66
13	Drilling.	0.2Min	INR25/Hr	INR 0.10
14	Tapping.	0.30Min	INR 25/Hr	INR 0.15

The production cost per piece of the contact is Rs.6.07. Add hidden cost 5% & the net cost =Rs.6.37/Pc.

It is well understood that the process is non productive and gives poor quality. Cannot be used further to satisfy customer's higher quantity demands. Hence search of better manufacturing technique is must.

Research content Phase -2: This mainly covered study of probable alternative techniques through industrial surveys and literature review this is as per Table 3.

Table 3 Industrial Survey findings

No	Name	Advantages	Limitations
1	Die forge.	Better Material Utilization	Further finishing by machining required.
2	Powder Metallurgy	Finish parts. No machining	Higher set up and tool cost.
3	Press Tool .	Productive for low thickness parts.	Poor tool life. Higher tool maintenance cost.
4	Extrusion	Higher productivity. More Yield..	Higher set up cost. Usable only for long sections.

The above table concludes that for the said contacts it is preferable to select forward hot extrusion as metal forming technique to extrude T-sectional long strips with additional allowance of 0.25 mm per

face and further to use drawing technique to remove this allowance to get finish section dimensions. Drawing being cold working operation gives better surface finish and closer tolerances as per drawing in figure 1.

Research content Phase3: Practical implication of Hydraulic forward extrusion and drawing operations. These are very high productive techniques hence decided not to make any investment in these set ups but to outsource from reliable sources with available spare capacity. The principle of outsourcing is no investment needed in set up. We buy only technology and overhead cost saved.

2.2 Extrusion Die Development

One important step of extrusion die design & development activity is explained in following Bill of Material Table 4.

Table 4 Extrusion die design & development activity

No	Part	R/ M	Heat treatment done if any
1	Sub bolster	H13	Std part of press
2	Bolster	H11	Std part.
3	Backer	H11	Not required.
4	Die	H13	Quench and temper.
5	Die ring.	H11	Not required.
6	Die slide	H11	Not required.
7	Dummy block	H11	Not required.

The manufacturing of various die components is done by using Machines such as ram turret milling, spark erosion and wire cutting ,as per need lapping is also done manually. The half sectional view of two cavity Extrusion die is as shown in the Figure 2.

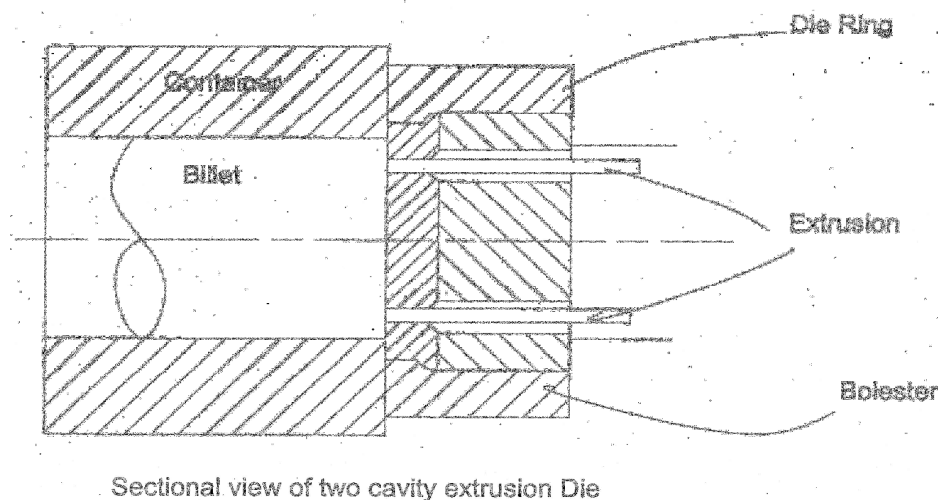


Figure 2 Sectional view of two cavity extrusion Die

2.3 Extrusion Press Capacity estimation

Other key area is finding extrusion press capacity. This works out to 250 MT from standard practice based on brass max shear stress and extrusion Ration. This is cross verified by comparing capacity required for extrusion of similar section components. For initial production runs single cavity die design is used so that further two cavity die can be used after freezing die dimensions. The shrinkage allowance provided is 2% on all linear cross sectional dimensions for brass. The weight of input and output of brass remains constant in the process. Input weight = Output weight .On these basis output extrusion sectional length is calculated. Extrusion ram speed is output speed of the cross section extruded. This is 1200 mm / Min. & slitting saw width is decided 1 mm. By slitting we cut extruded sections to suitable lengths. In this case this is 1000 mm. After extrusion these long sections are further drawn so as to get cross sectional finish dimensions. After slitting we have to do three machining operations viz.1 Side slot milling by using SPM & HSS side and face cutter as cutting tool.2 Drilling by using SPM .3Tapping by tapping machine. For design and development of These

SPMS we have made a two engineers team dedicated to work with SPM developing vendor. The prime objective of developing special purpose machines is to reduce manual handling of work piece to reduce labor fatigue and fool proofing of machining operations. The production sequence and cost estimation by using this technique as shown in the table Table 5.

Table 5 Production sequence and cost estimation

No	Name of the operation	Machine or set up used	Estimate Base .	Per contact Cost.
1	Raw material cost	Market Rate	INR 315/Kg	INR.3.15
2	Billet casting Pouring.	Melting +Pouring into molds	INR 20/kg	12 Kg Billet wt=INR 240/Billet =INR 0.25
3	Heating of the billet	Oil/gas fired furnace	INR 10/kg	INR 120/billet= INR 0.12
4	Die Design, development	Tool room machines	48 Tool room Hrs. INR 250/Hr	INR 12,000/- is effective of INR 0.12
5	Loading the die on extrusion press		Manual(30 Two persons)	INR. 31.25 /Loading INR=0.05
6	Extrusion and drawing	Hydraulic extrusion press.	INR/Hr 280	Ram Speed 1200 mm. /Min=INR 0.05
7	Slitting the extrusion to suitable length	Slitting saw	INR 8/Hr Time 0.5 min	INR 0=06
8	Inspection	Manual	INR 30/Hr	INR0.02/Pc
9	Milling	SPM	INR80/Hr	INR0.22/Pc
10	Drilling	SPM	INR12/Hr.	INR0.05/Pc
11	Tapping	SPM	INR12/Hr	INR0.04/Pc

The total number of operations is 11 Nos.

2.4. Commercial impact of the Research Cost per piece of the contacts as per advance manufacturing technology is INR 4.13 /Pc add 5% in this for hidden cost.=INR 4.33 /Pc. Cost saved per contact INR 2.04.Cost saved /Month=INR 612000 Qty. Basis 300000 is average
Annual savings= INR 612000 assuming ten month assuming ten months business per year.

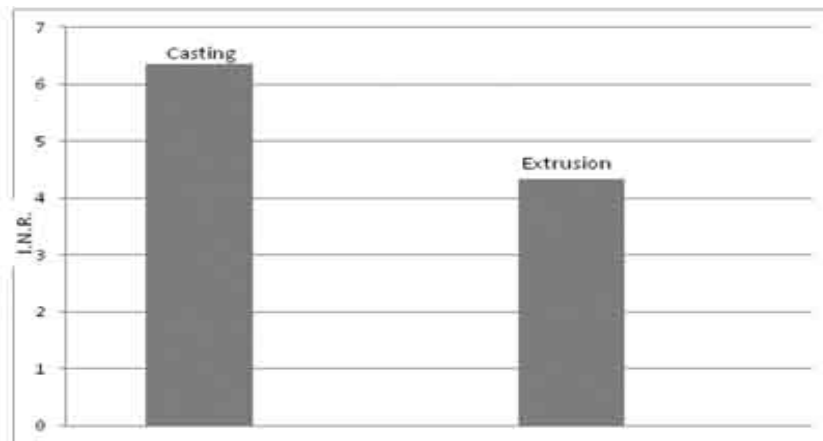


Figure 3 Cost per contact comparison

2.5 Quality improvement

Initial cumulative rejection stages =47%. Rejection after using advance manufacturing technique average -7 %.Refer graph plotted from for three months field rejection data for both casting and extrusion process .Figure 4 as below. The Graph shows reduction in percentage rejection from 47 to average 7 percentage for 3 months.

Figure 4 indicates process rejection.

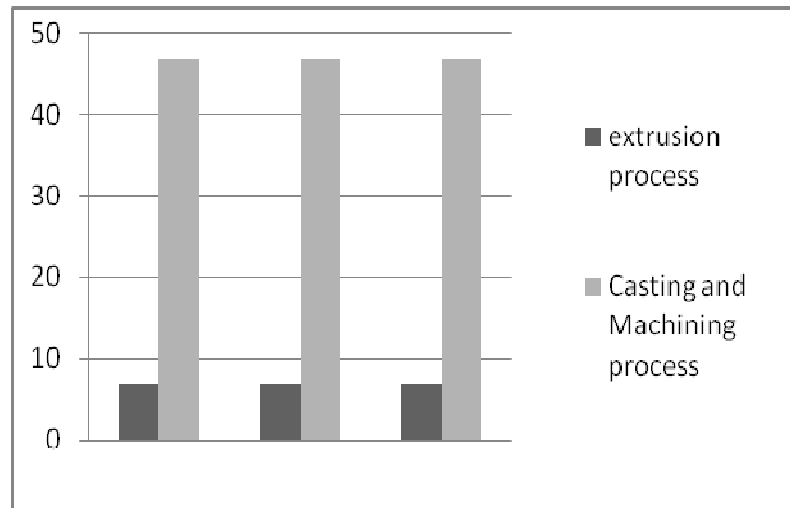


Figure 4 Reduction in percentage rejection

2.6 Lead Time Reduction

Lead time reduced from 15 days to 7 days a batch of 100 000 Nos.

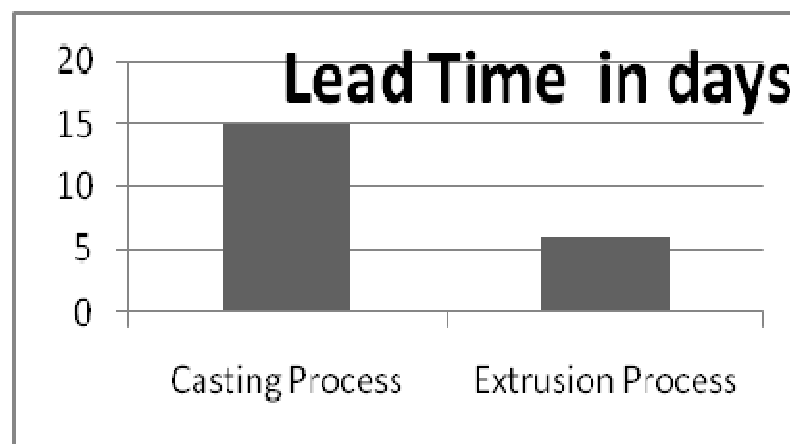


Figure 5 Lead Time Reduction

Figure 5 clearly indicates reduction in lead time.

2.7 Overall advantages

The overall advantages are lead time reduction and productivity and quality enhancement. Assembly function runs in time. Customer expectations satisfied in time.

2.8 Results and Discussions

The estimated cost as per casting and machining process used initially was INR 6.21 per piece of electrical contact and after implementation of extrusion technique, the cost reduced to INR 4.10 i.e. there is net cost reduction of INR 2.11 per piece or say around 33%. Any saving done due to use of advanced manufacturing technique directly contributes to the net profit of the product. Due to high cost competition in the market, this cost reduction benefit is passed on to the customer to make product more competent in the global market. If we look at cycle time required for completing a batch of contacts we see that previously lead time required was 15 days. This is now reduced to 7 days. In the long run manufacturer can think of reducing wip inventory. This is again long term benefit to the organization. This research has extended product life cycle of the said product which was landed into trouble at mass production stage itself.

III. CONCLUSIONS

Engineering manufacturing houses engaged in the design and development of products and selling their products in high competitive market must go for study and practicing use of advance

manufacturing techniques suitable to all class of components with priority for A class components not at latter stage but at the design, development and prototyping stage of the components. This derives long term benefits to the organization which mainly includes higher productivity, consistent quality and more competitive prices for sales of the product. By this organization is in position to sell the product at lower prices and capture more and more market at global levels. The cost reduction is also achieved because of material waste reduction due to utilization of advanced manufacturing techniques. Due to higher productivity the return on investment also starts at initial phase of mass production. This is long term economic advantage to manufacturing organizations and they should consistently strive for establishing a system of appropriate manufacturing technique selection through team working with special focus on continuous assessment and updated knowledge of advanced manufacturing techniques.

ACKNOWLEDGEMENTS

We are heartily thankful to Vikrant Industries, Jamnagar, Gujarat, INDIA for allowing us to do field work for solving their bottle neck in Electric Contact manufacturing and providing necessary resources and setup for performing necessary research and trials.

REFERENCES

- [1] Rao V. Dukkipatty, Pradip K. Ray. Process design for economy and reliability Publication. Publication – New Age International –First Edition I.S.B.N.-978-81-224-2661.
- [2] ASM Handbook Volume-14,Forming and forging ASM International –The materials information society- ISBN-87170-007-7(V1) SAN-204-7586 Page. No-255- 259
- [3] Richard W. Heine, Carl Loper, Philip C. Rosenthal Principles Of Metal Casting , Publication-- Tata Mc. Grawhill,ISBN-13,978-0-07-099348-.
- [4] Philipps Ostwald, Jairo Munz, Manufacturing Processes and Systems by Pub. – Wiley, India Pvt. Ltd. ISBN-978-81-265-1894-
- [5]Fundamentals of Tool Design – American Society of Tool and Manufacturing Engineers, A Publication of ASTME Manufacturing Engineering Series, Pub. Prentice Hall Inc. New Jersey.)
- [6] Process and design for manufacturing Sheriff Diwakil, Narinder Kumar Lijhara for overseas press India (P) Ltd.ISBN 81-88689-70
- [7] Fundamentals of metal forming processes By B.L. Juneja –New edge international Publishers ISBN :978-81-224-3089-9.Pg no 274-303.

Authors

Virajit A. Gundale is presently working as the Professor & Academic In-charge at Sharad Institute of Technology College of Engineering, Yadav Dist. Kolhapur, India. He obtained his B.E. in Mechanical Engineering from Shivaji University and M.S. in Manufacturing Management from BITS, Pilani. He obtained his Ph.D. in Manufacturing Technology from UNEM, Costa Rica in 2010. His total experience including Teaching and Industry spans more than 11 years. He is also a fellow of the International Institute of Mechanical Engineers, South Africa.



Vidyardhar M. Dandge is presently working as the Assistant Professor at Dr.J.J.Magdum College Of Engineering Jaysingpur, Dist-Kolhapur, India. He obtained his B.E. in Mechanical Engineering from Karnataka University and M.E (Production) from Shivaji University. His total experience including Teaching and Industry spans more than 25 years.



CONTROL AND PERFORMANCE OF A CASCADED H-BRIDGE MLI AS STATCOM

M. Vishnu Prasad and K. Surya Suresh

Department of Electrical & Electronics Engineering, SVIET, Nandamuru, AP, India

ABSTRACT

This paper presents a three-phase cascaded H-Bridge MLI as Static synchronous Compensator (STATCOM) intended for installation on industrial and utility power distribution systems. It proposes a control technique that devotes itself to meeting the demand of reactive power. To implement a separate control for the three phase dc-link voltages, the average active power in each phase can also be adjusted to a target value determined by the dc-link voltage control loop. Then, by forcing the converter neutral voltage to be equal to the counterpart of the equivalent power supply, the STATCOM can be decoupled into three single-phase systems and the line-to-neutral voltage of the equivalent power supply can be used as the input voltage to the corresponding phase leg. Accordingly dc-link voltage maintaining can be simultaneously achieved under unbalanced conditions.

KEYWORDS— STATCOM, Multilevel inverter, Reactive power, FACTS.

I. INTRODUCTION

An AC power system is a complex network of synchronous generators, transmission lines and loads. The transmission lines can be represented mostly as reactive networks composed of series inductors and shunt capacitors. The total series inductance, which is proportional to the length of the line, determines primarily the maximum transmissible power at a given voltage. The shunt capacitance influences the voltage profile and thereby the power transmission along the line.

The transmitted power over a given line is determined by the line impedance, the magnitude of voltage and phase angle between the end voltages, the basic operating requirements of an AC power system are that the synchronous generators must remain in synchronism and the voltage must kept close to their rated values.

In the late 1980s, the Electric Power Research Institute (EPRI) in the USA formulated the vision of the Flexible AC transmission System (FACTS) in which various Power electronics based controllers regulate Power flow and transmission voltage through rapid control action, mitigate dynamic disturbances. FACTS devices involve the applications of high power electronics in AC transmission networks enables fast and reliable control of power flows and voltages.

Reactive Power compensation can be obtained by Series VAR compensation and Shunt VAR compensation. Series compensation modifies the transmission or distribution system parameters, while shunt compensation changes the equivalent impedance of the load. Traditionally, rotating synchronous condensers and fixed or mechanically switched capacitors or inductors have been used for reactive power compensation. However, in recent years, static VAR compensators employing thyristor switched capacitors and thyristor controlled reactors to provide or absorb the required reactive power have been developed [1][3].

The FACTS is a concept based on power-electronic controllers, which enhance the value of transmission networks by increasing the use of their capacity As these controllers operate very fast, they enlarge the safe operating limits of a transmission system without risking stability. FACTS

controllers can be classified as Series controllers, Shunt controllers, Combination of Series - Shunt controllers and Combination of Series- Series controllers.

The STATCOM is a shunt-connected reactive-power compensation device that is capable of generating or absorbing reactive power and in which the output can be varied to control the specific parameters of an electric power system [5].

Multilevel inverter has a series of advantages over two-level converter, such as its output waveforms are more similar to the object modulation waves (sinusoidal waves), less skip and harmonics of output voltages, fewer switching losses [12].

II. REACTIVE POWER COMPENSATION TECHNOLOGIES

VAR compensation is defined as the management of reactive power to improve the performance of ac power systems. The concept of VAR compensation embraces a wide and diverse field of both system and customer problems, especially related with power quality issues, since most of power quality problems can be attenuated or solved with an adequate control of reactive power. In general, the problem of reactive power compensation is viewed from two aspects: load compensation and voltage support. In load compensation the objectives are to increase the value of the system power factor, to balance the real power drawn from the ac supply, compensate voltage regulation and to eliminate current harmonic components produced by large and fluctuating nonlinear industrial loads. Voltage support is generally required to reduce voltage fluctuation at a given terminal of a transmission line. Reactive power compensation in transmission systems also improves the stability of the ac system by increasing the maximum active power that can be transmitted. It also helps to maintain a substantially flat voltage profile at all levels of power transmission; it improves HVDC (High Voltage Direct Current) conversion terminal performance, increases transmission efficiency, controls steady-state and temporary over-voltages, and can avoid disastrous blackouts[16].

Traditionally, rotating synchronous condensers and fixed or mechanically switched capacitors or inductors have been used for reactive power compensation. However, in recent years, static VAR compensators employing thyristor switched capacitors and thyristor controlled reactors to provide or absorb the required reactive power have been developed. Also, the use of self-commutated PWM converters with an appropriate control scheme permits the implementation of static compensators capable of generating or absorbing reactive current components [1] [2] with a time response faster than the fundamental power network cycle.

III. FACTS CONTROLLERS

The rapid growth in electrical energy use, combined with the demand for low cost energy, has gradually led to the development of generation sites remotely located from the load centers. In particular, the remote generating stations include hydroelectric stations, which exploit sites with higher heads and significant water flows; fossil fuel stations, located close to coal mines; geothermal stations and tidal-power plants, which are site bound; and, sometimes, nuclear power plants purposely built distant from urban centers. The generation of bulk power at remote locations necessitates the use of transmission lines to connect generation sites to load centers. Furthermore, to enhance system reliability, multiple lines that connect load centers to several sources, interlink neighboring utilities, and build the needed levels of redundancy have gradually led to the evolution of complex interconnected electrical transmission networks. These networks now exist on all continents [7].

The FACTS is a concept based on power-electronic controllers, which enhance the value of transmission networks by increasing the use of their capacity. As these controllers operate very fast, they enlarge the safe operating limits of a transmission system without risking stability. Needless to say, the era of the FACTS was triggered by the development of new solid-state electrical switching devices [4]. Gradually, the use of the FACTS has given rise to new controllable systems.

Today, it is expected that within the operating constraints of the current-carrying thermal limits of conductors, the voltage limits of electrical insulating devices, and the structural limits of the supporting infrastructure, an operator should be able to control power flows on lines to secure the highest safety margin as well as transmit electrical power at a minimum of operating cost. Doing so constitutes the increased value of transmission assets.

Types of FACTS Controllers:

In general FACTS controllers can be divided into the following four categories:

a. Series Controllers

In principle all the series controllers inject voltage in series with the line. Series connected controller impacts the driving voltage and hence, the current and power flow directly. Static Synchronous Series Compensator (SSSC), Thyristor Controlled Series Compensator (TCSC) etc. are the examples of series controllers.

b. Shunt Controllers

All shunt controllers inject current into the system at the point of connection. The shunt controller is like a current source, which draws/injects current from/into the line. Static Synchronous Compensator (SSC), Static Synchronous Generator (SSG), Thyristor Controlled Reactor (TCR) etc are the examples of shunt controllers.

c. Combined Series-Shunt Controllers

This could be a combination of separate shunt and series controllers, which are controlled in a coordinated manner. Combined shunt and series controllers inject current into the system with the shunt part of the controller and voltage in series in the line with the series part of the controller. Unified Power Flow Controller (UPFC) and Thyristor Controlled Phase Shifting Transformer (TCPST) are the examples of shunt series controllers.

d. Combined Series-Series Controllers

This could be a combination of separate series controllers, which are controlled in a coordinated manner, in a multi-line transmission system or it could be a unified controller, in which series controller provides independent series reactive compensation for each line but also transfer real power among the line via the power link.

IV. STATCOM

The STATCOM (or SSC) is a shunt-connected reactive-power compensation device that is capable of generating or absorbing reactive power and in which the output can be varied to control the specific parameters of an electric power system. It is in general a solid-state switching converter capable of generating or absorbing independently controllable real and reactive power at its output terminals when it is fed from an energy source or energy-storage device at its input terminals. Specifically, the STATCOM considered in this project is a voltage-source converter that, from a given input of dc voltage, produces a set of 3-phase ac-output voltages, each in phase with and coupled to the corresponding ac system voltage through a relatively small reactance. The dc voltage is provided by an energy-storage capacitor[5][6].

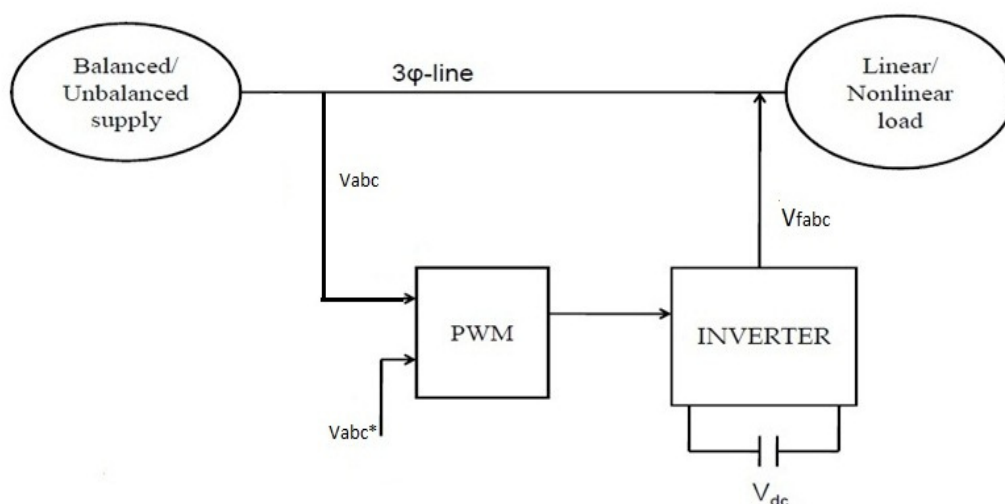


Fig.1 STATCOM Principal Diagram

A STATCOM can improve power-system performance in such areas as the following:

The dynamic voltage control in transmission and distribution systems, The power-oscillation damping in power-transmission systems, The transient stability, The voltage flicker control and The control of

not only reactive power but also (if needed) active power in the connected line, requiring a dc energy source.

A STATCOM is analogous to an ideal synchronous machine, which generates a balanced set of three sinusoidal voltages—at the fundamental frequency—with controllable amplitude and phase angle. This ideal machine has no inertia, is practically instantaneous, does not significantly alter the existing system impedance, and can internally generate reactive (both capacitive and inductive) power.

4.1 Principle of Operation

A STATCOM is a controlled reactive-power source. It provides the desired reactive-power generation and absorption entirely by means of electronic processing of the voltage and current waveforms in a voltage-source converter (VSC) [7]. A STATCOM is seen as an adjustable voltage source behind a reactance—meaning that capacitor banks and shunt reactors are not needed for reactive-power generation and absorption, thereby giving a STATCOM a compact design, or small footprint, as well as low noise and low magnetic impact.

The exchange of reactive power between the converter and the ac system can be controlled by varying the amplitude of the 3-phase output voltage, V_s , of the converter. That is, if the amplitude of the output voltage is increased above that of the utility bus voltage, V_t , then a current flows through the reactance from the converter to the ac system and the converter generates capacitive-reactive power for the ac system. If the amplitude of the output voltage is decreased below the utility bus voltage, then the current flows from the ac system to the converter and the converter absorbs inductive-reactive power from the ac system. If the output voltage equals the ac system voltage, the reactive-power exchange becomes zero, in which case the STATCOM is said to be in a floating state.

Adjusting the phase shift between the converter-output voltage and the ac system voltage can similarly control real-power exchange between the converter and the ac system. In other words, the converter can supply real power to the ac system from its dc energy storage if the converter-output voltage is made to lead the ac-system voltage. On the other hand, it can absorb real power from the ac system for the dc system if its voltage lags behind the ac-system voltage.

A STATCOM provides the desired reactive power by exchanging the instantaneous reactive power among the phases of the ac system. The mechanism by which the converter internally generates and/or absorbs the reactive power can be understood by considering the relationship between the output and input powers of the converter. The converter switches connect the dc-input circuit directly to the ac-output circuit. Thus the net instantaneous power at the ac output terminals must always be equal to the net instantaneous power at the dc-input terminals (neglecting losses)[8][9].

Although reactive power is generated internally by the action of converter switches, a dc capacitor must still be connected across the input terminals of the converter. The primary need for the capacitor is to provide a circulating-current path as well as a voltage source. The magnitude of the capacitor is chosen so that the dc voltage across its terminals remains fairly constant to prevent it from contributing to the ripples in the dc current.

The VSC-output voltage is in the form of a staircase wave into which smooth sinusoidal current from the ac system is drawn, resulting in slight fluctuations in the output power of the converter [10]. However, to not violate the instantaneous power-equality constraint at its input and output terminals, the converter must draw a fluctuating current from its dc source. Depending on the converter configuration employed, it is possible to calculate the minimum capacitance required to meet the system requirements, such as ripple limits on the dc voltage and the rated-reactive power support needed by the ac system.

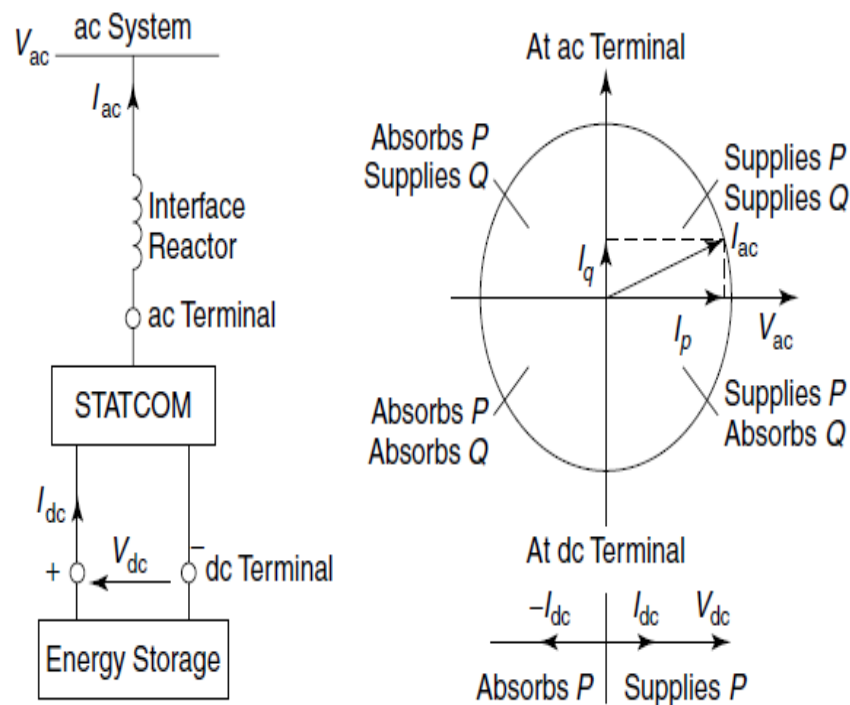


Fig.2 The power exchange between the STATCOM and the ac system

4.2 A Multilevel VSC-Based STATCOM

The harmonic contamination of the power-system network by the addition of STATCOM into the power system can be reduced by employing multilevel VSC configurations. The multilevel converters usually synthesize a staircase-type voltage wave from several levels of dc-voltage sources (typically capacitor-voltage sources). The multilevel VSC schemes studied and tested so far include the diode clamp, the flying capacitor, and the cascaded, separate dc-source converter types. Many separate dc-source converter topologies are suggested in refs. Multilevel converters can reach high voltages and reduce harmonic distortion because of their structure. To increase the voltage rating, many single-phase full-bridge converters (FBCs) can be connected in series, automatically leading to a desirable reduction of harmonic distortion. However, the need to balance capacitor voltages, the complexity of switching, and the size of the capacitors all limit the number of levels that can be practically employed [16][17].

Figure 3 shows the 3-phase star-connected arrangement of the separate dc-source, 3-level binary VSC commonly referred to as a BVSI. It consists of three single-phase FBCs, each with its own dc source, connected in series. However, the magnitude of each dc source is in binary proportion of V_{dc} , $2V_{dc}$, and $4V_{dc}$, where V_{dc} is chosen to get the desired fundamental ac-voltage output for a normalized 1-pu modulation index. The switches are turned on and off to generate a 15-step ac-voltage output over one fundamental cycle. In general, n -level BVSI would produce a $(2^{n+1}-1)$ -step ac-voltage output versus a $(2n+1)$ -step output generated by a conventional n -level, separate dc-source VSC configuration.

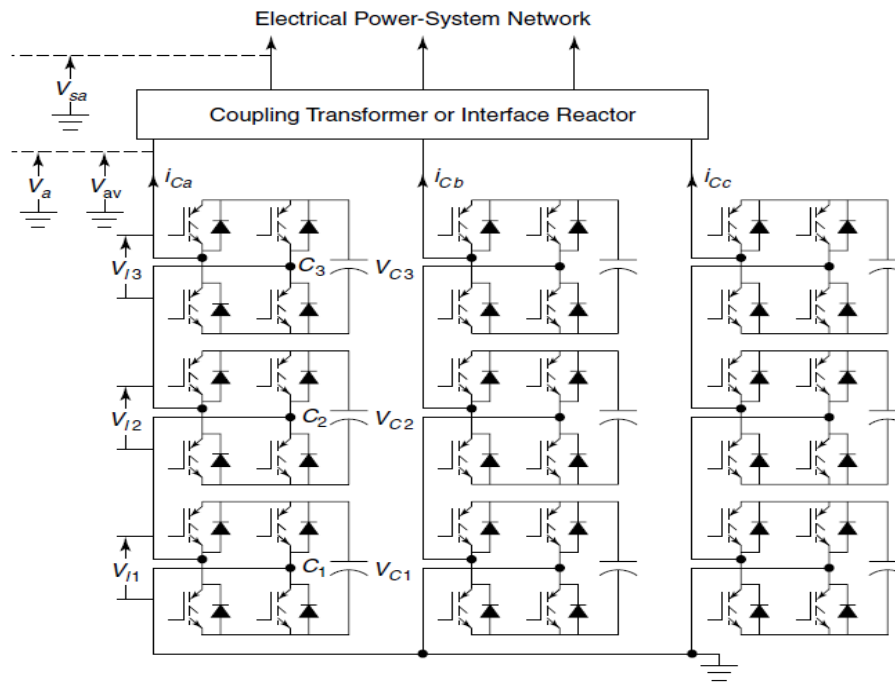


Fig 3 The 3-phase, star-connected 3-level VSI

The multilevel converters usually synthesize a staircase-type voltage wave from several levels of dc-voltage sources (typically capacitor-voltage sources). The multilevel VSC schemes studied and tested so far include the diode clamp, the flying capacitor, and the cascaded, separate dc-source converter types. Many separate dc-source converter topologies are suggested in refs. Multilevel converters can reach high voltages and reduce harmonic distortion because of their structure. To increase the voltage rating, many single-phase full-bridge converters (FBCs) can be connected in series, automatically leading to a desirable reduction of harmonic distortion. However, the need to balance capacitor voltages, the complexity of switching, and the size of the capacitors all limit the number of levels that can be practically employed [7].

V. CONTROL STRATEGY

Active- and Reactive-Power Controls

The control strategy that is proposed in this paper regulates a generic single-phase cascaded H-bridge multilevel converter, which is composed of N H-bridges that are connected in series. The connection of the STATCOM device to the power grid is made using a coupling inductance. Assume that the grid voltage and output-current expressions are equal to the following:

$$V_{\text{Grid}} = \sqrt{2} V_{\text{Grid}} \cdot \cos(\omega t) \quad (1)$$

$$i_f = \sqrt{2} I_f \cdot \cos(\omega t - \phi) \quad (2)$$

The expressions of the active and reactive powers that are supplied by the STATCOM to the grid are the following

$$P_{\text{TOTAL}} = V_{\text{Grid}} \cdot I_f \cdot \cos(\phi) \quad (3)$$

$$Q_{\text{TOTAL}} = V_{\text{Grid}} \cdot I_f \cdot \sin(\phi) \quad (4)$$

Where ϕ is the angle between the grid voltage and the current that is injected by the STATCOM. The active power is proportional to the current component which is in phase with the grid voltage (active component), and the reactive power is proportional to the current component which is orthogonal to the grid voltage (reactive component). Therefore, for active-power regulation, the active component of the current is changed, and on the other hand, for reactive-power regulation, the reactive component is modified. From these two current components, the instantaneous current reference could be generated.

The reactive-power reference is obtained from the dc-bus voltage control loops. If the dc-bus voltage of one of the shunt-connected H-bridges is lower than the reference voltage, it means that the reactive power should be provided to this H-bridge; therefore, the capacitor would take energy, which would increase the voltage level. On the other hand, if the dc-bus voltage is higher than the voltage reference, energy should be taken out of the bus capacitor. In this way, the output of each dc-bus voltage regulator is a particular reactive-power reference for each H-bridge.

VI. SIMULATION RESULTS

To verify the performance of the proposed cascade Three phase five level inverter is implemented as STATCOM with phase shifted bipolar sinusoidal PWM as control algorithm. Simulation have been carried out using Matlab–Simulink.

In this paper, a simulation block set in Simulink/matlab was implemented as shown in fig 4 and the results are presented. In this model initially load1 (50KW, 15 KVAR) is supplied with 415V. After a period of 2 cycles another load2 (30KW, 10KVAR) is added with the help of circuit breaker. A five level inverter is connected as STATCOM. The control signal for this STATCOM taken from a closed loop POD technique. In this technique load voltage is always compared with reference voltage, whenever a deviation in the load voltage an error signal generated in the form of sine wave and this sine wave is used as reference signal in POD technique. Hence this control system produces necessary control signals for STATCOM to compensate the load voltage and reactive power requirement of the system.

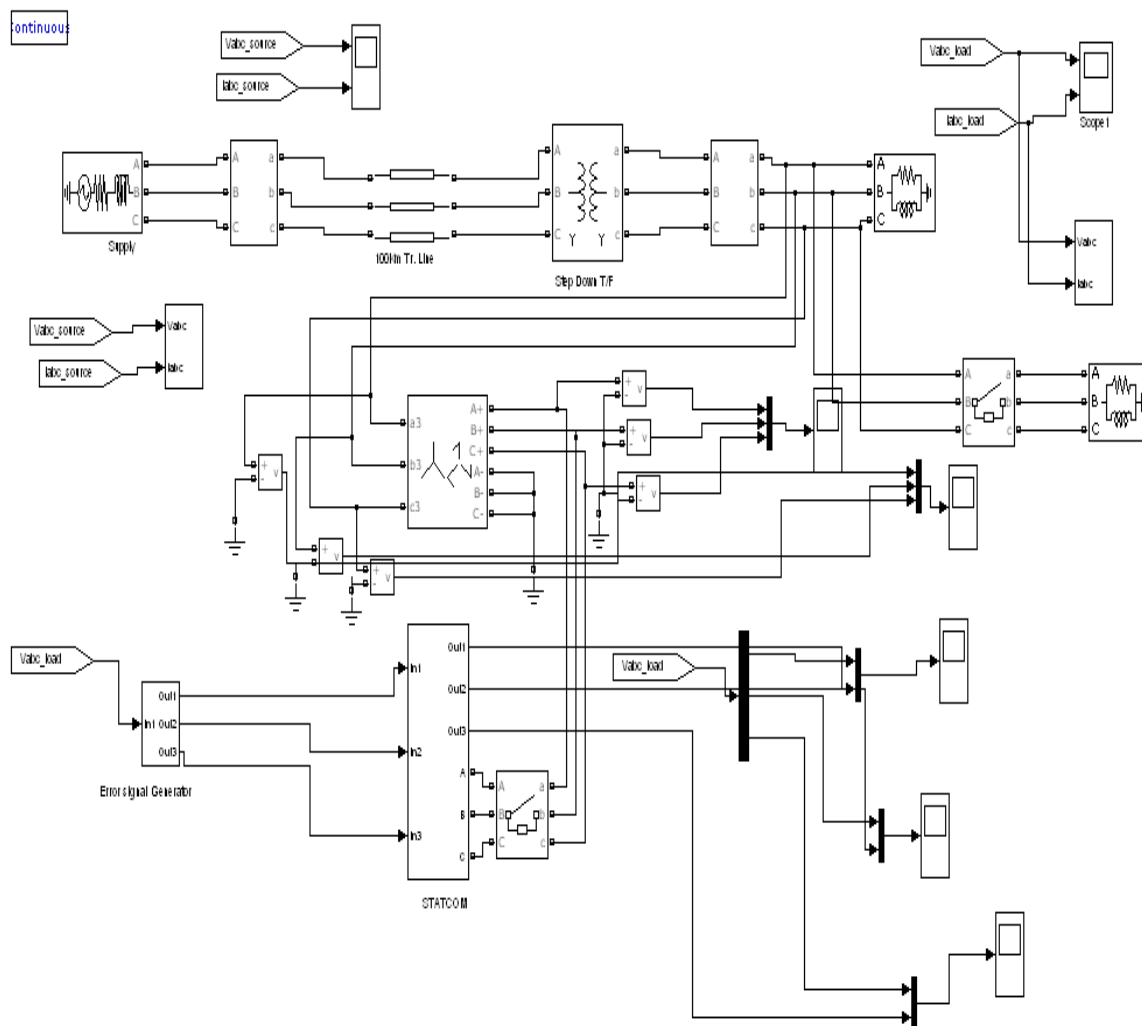


Fig.4 Simulation Diagram of MLI as STATCOM 1

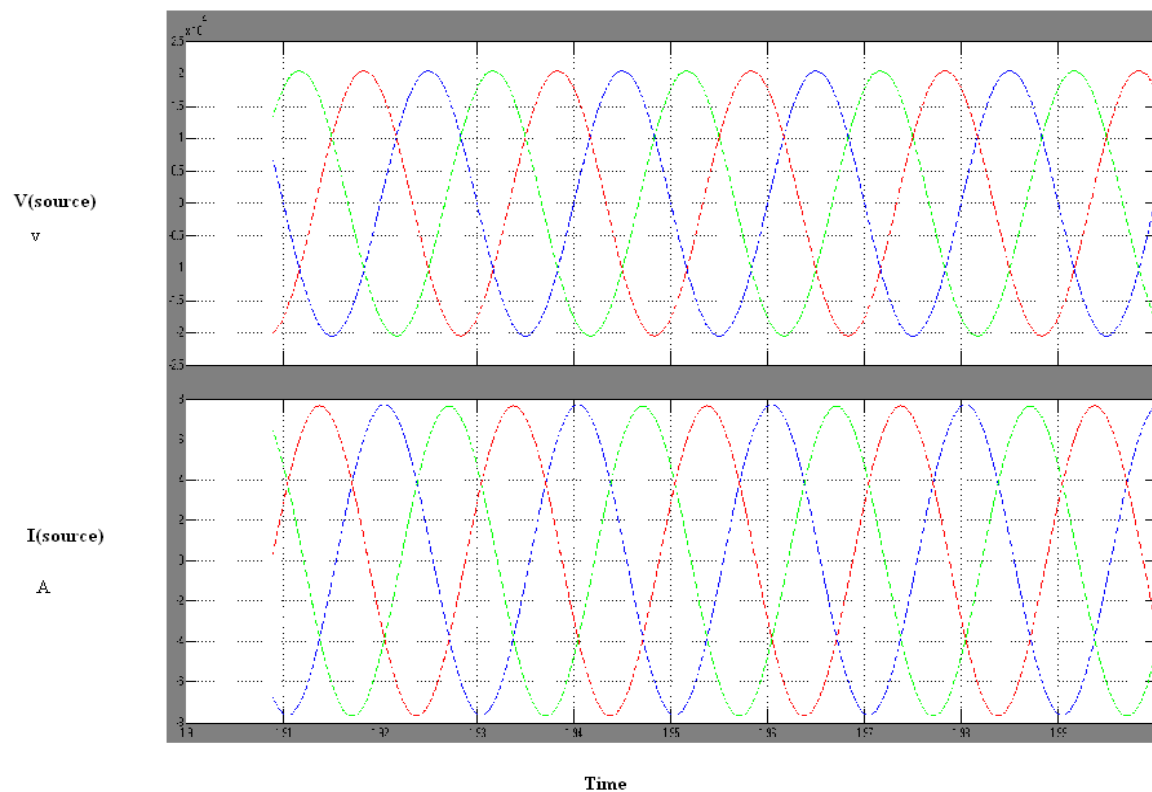


Fig.5 Source Voltage and Source Current 1

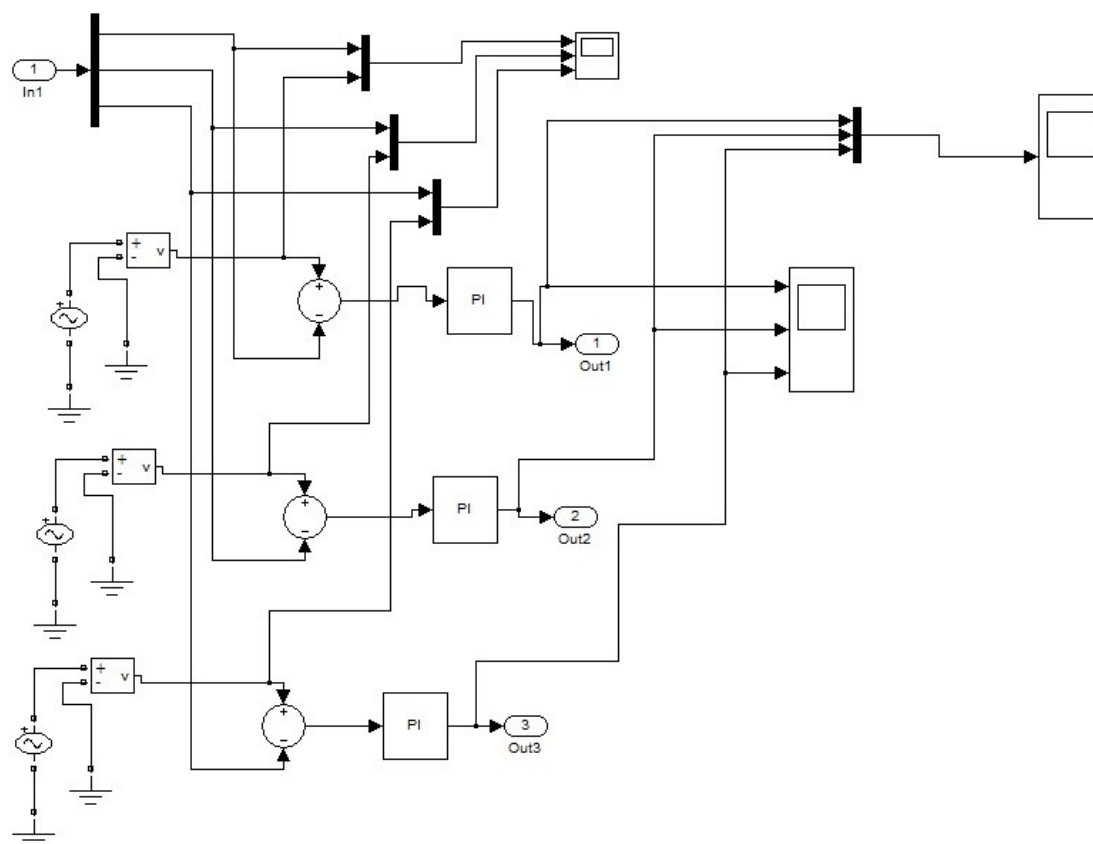


Fig. 6 Error voltage generator 1

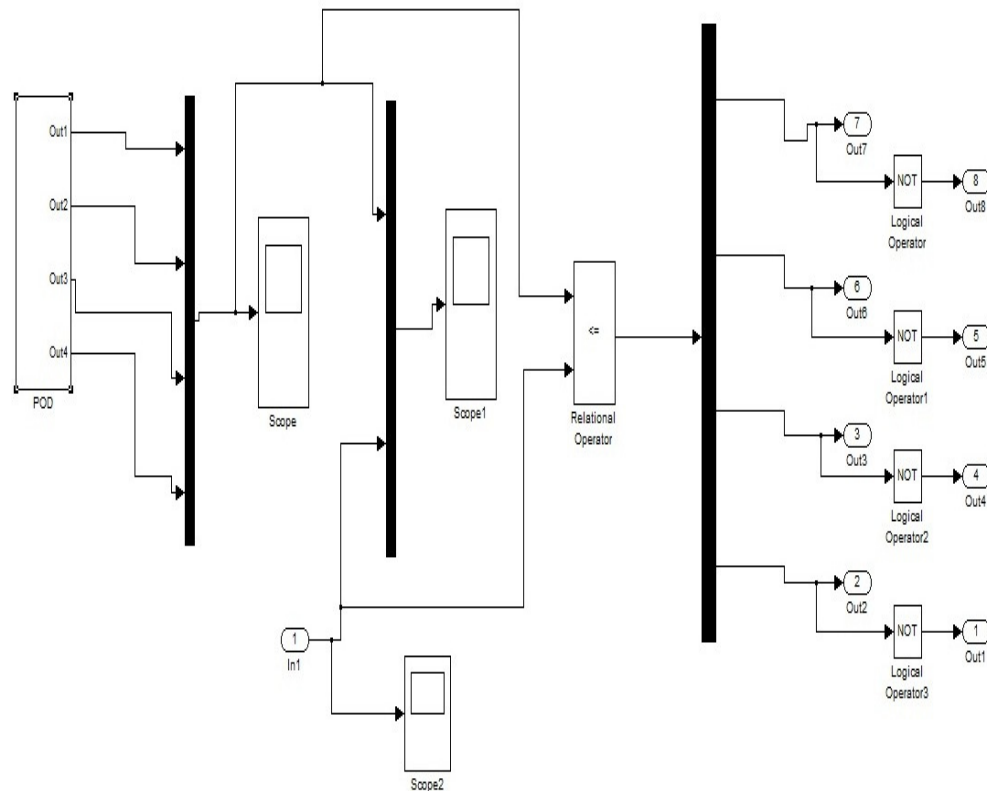


Fig. 7 Pulse Generation for STATCOM 1

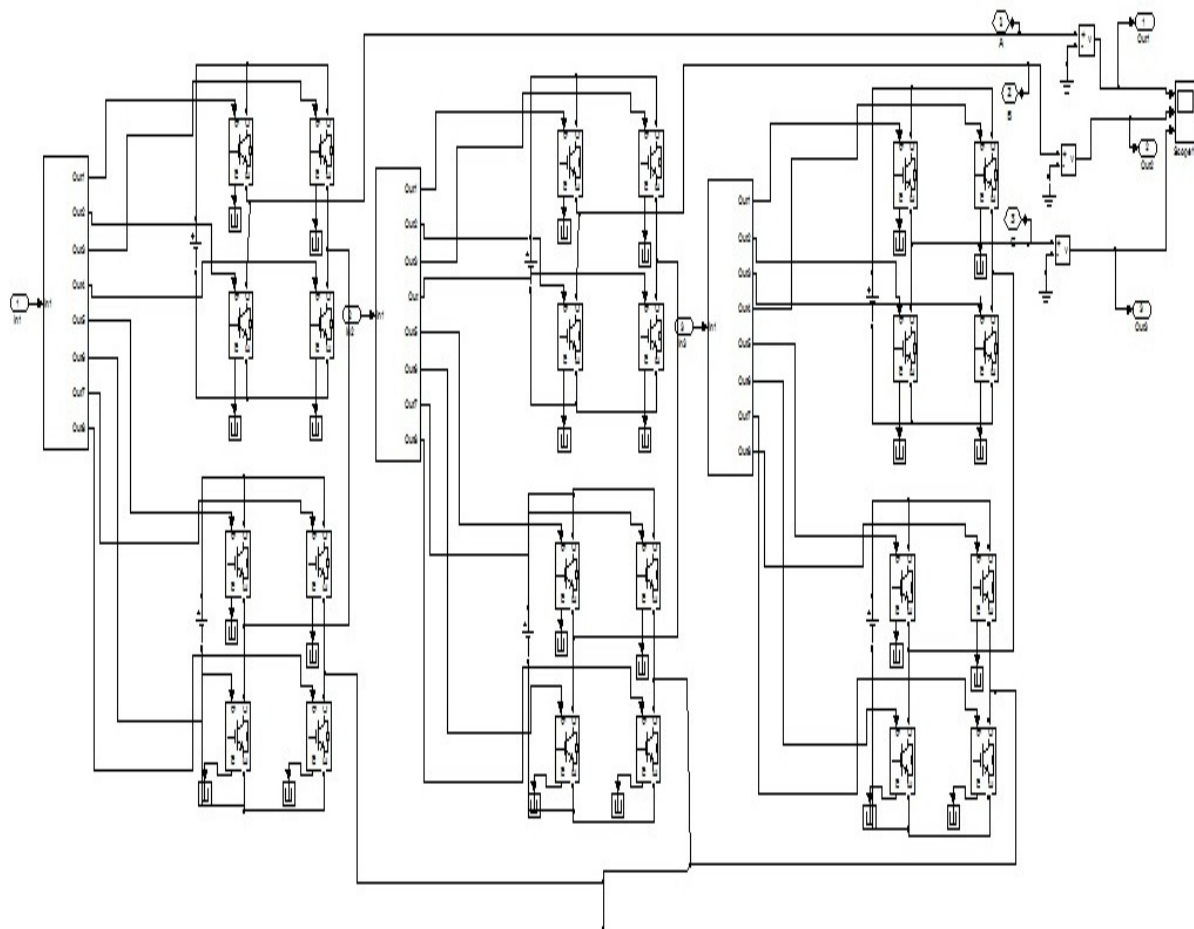


Fig. 8 MLI using as STATCOM 1

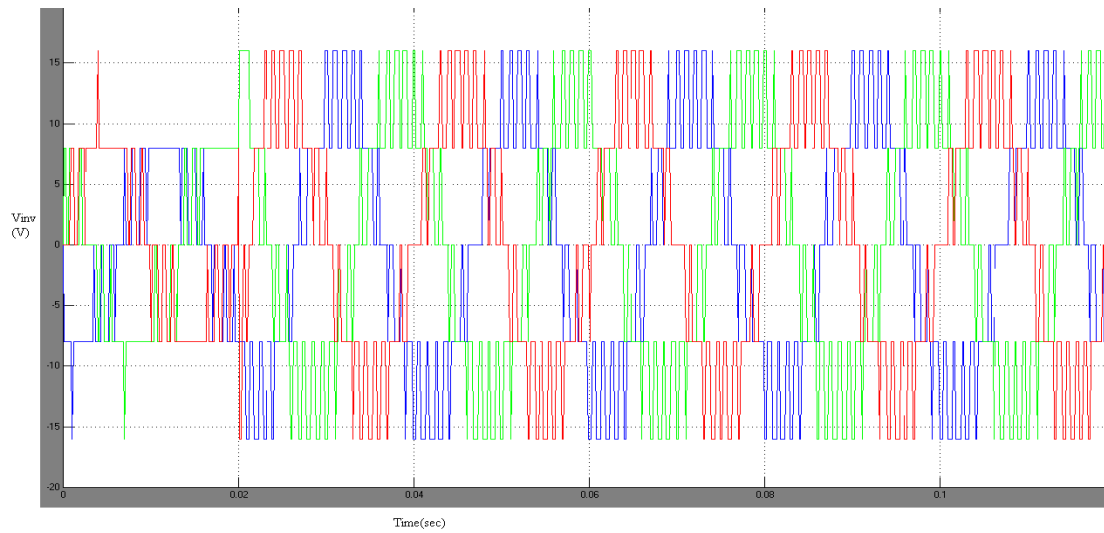


Fig. 9 STATCOM output voltage 1

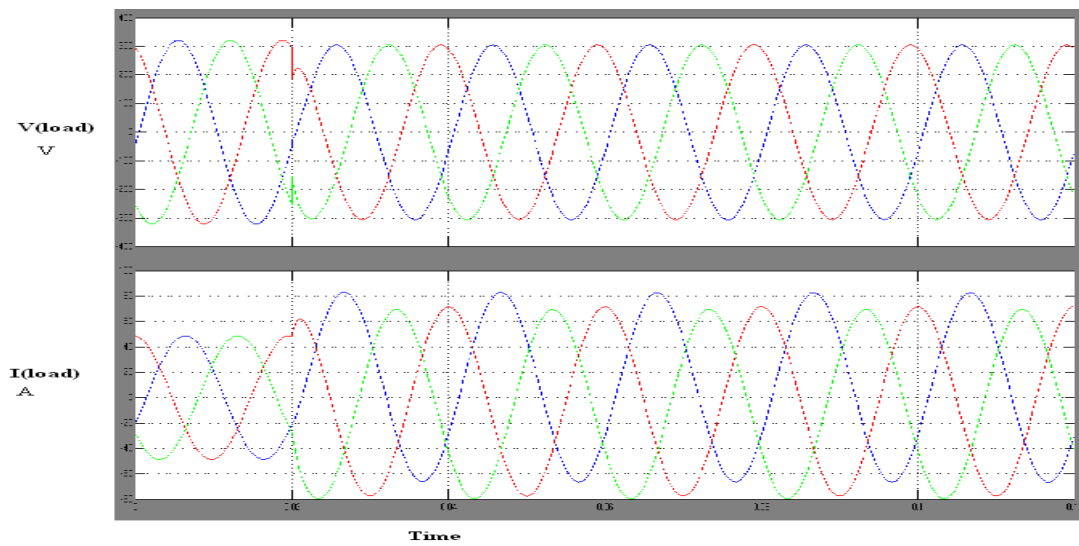


Fig10 Load Voltage and Current 1

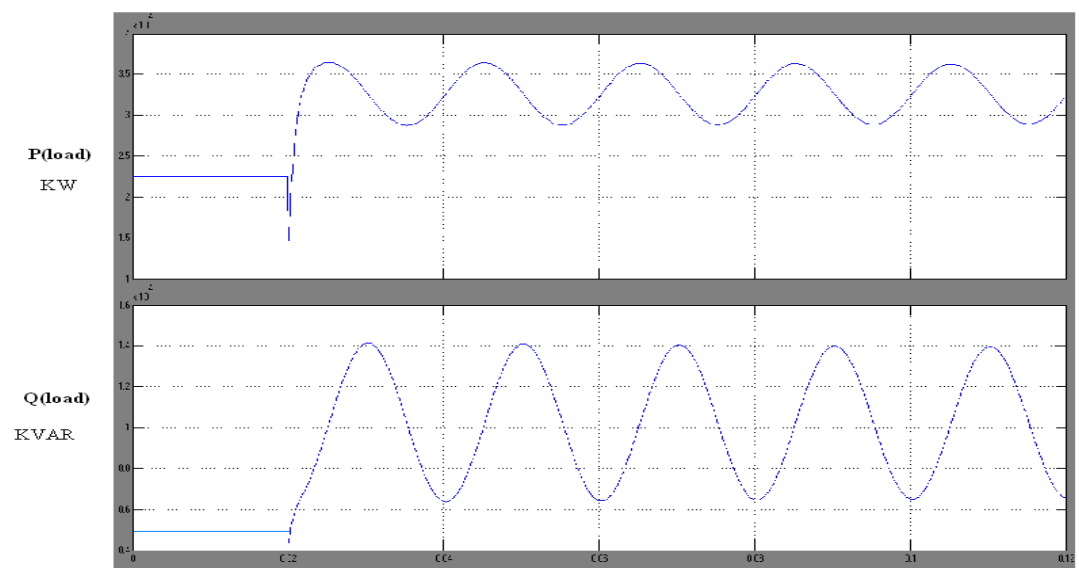


Fig. 11 Load Active and reactive Powers 1

Fig 5 shows the balanced voltage and current supplying from the source. In fig 6 error signal generator has been shown. Fig 7 represents pulse generation for STATCOM based on the error signal. Fig 10 indicates the output voltage which clearly shows even after adding a load after two cycles it maintains a constant voltage where as a change in output current Fig 11 shows load active and reactive power when MLI acting as STATCOM.

VII. CONCLUSION

In this paper cascade five level inverter is implemented as STATCOM with star configuration for medium-voltage applications. The control algorithm is based on the “POD” type Multicarrier PWM. The control arrangement results from giving priority to the voltage-balancing control and reactive power balancing control. This model is mainly concentrated on unbalanced and continuously varying loads. Whenever there is a change in load this control arrangement of STATCOM able to balance the voltage and compensate the required reactive power. A 415V supply is given to the varying unbalanced load and a cascade number of $N = 2$ have verified the ability and effectiveness of the Reactive Power compensation.

REFERENCES

- [1] Hirofumi Akagi, Fellow, IEEE, Shigenori Inoue, Member, IEEE, and Tsurugi Yoshii “Control and Performance of a Transformerless Cascade PWM STATCOM With Star Configuration” IEEE Transactions on Industry Applications, Vol. 43, No. 4, July/August 2007
- [2] Ben-Sheng Chen and Yuan-Yih Hsu, Senior Member, “A Minimal Harmonic Controller for a STATCOM” IEEE Transactions On Industrial Electronics, Vol. 55, No. 2, February 2008
- [3] Qiang Song and Wenhua Liu, Member, IEEE “Control of a Cascade STATCOM With Star Configuration Under Unbalanced Conditions” IEEE Transactions On Power Electronics, Vol. 24, No. 1, January 2009
- [4] Chien-Hung Liu and Yuan-Yih Hsu, Senior Member, IEEE “Design of a Self-Tuning PI Controller for a STATCOM Using Particle Swarm Optimization” IEEE Transactions On Industrial Electronics, Vol. 57, No. 2, February 2010
- [5] Boštjan Blažič, Student Member, IEEE, and Igor Papić, Member, IEEE “Improved D-StatCom Control for Operation With Unbalanced Currents and Voltages” IEEE Transactions On Power Delivery, Vol. 21, No. 1, January 2006
- [6] Jon Andoni Barrena, Student Member, IEEE, Luis Marroyo, Member, IEEE, Miguel Ángel Rodríguez Vidal, Member, IEEE, and José Ramón Torrealday Apraiz “Individual Voltage Balancing Strategy for PWM Cascaded H-Bridge Converter-Based STATCOM” IEEE Transactions On Industrial Electronics, Vol. 55, No. 1, January 2008
- [7] Fang Zheng Peng, Senior Member, IEEE “A Generalized Multilevel Inverter Topology with Self Voltage Balancing” IEEE Transactions On Industry Applications, Vol. 37, No. 2, March/April 2001
- [8] Jianye Chen, shan song and Zhanji Wang “Analysis and Implement of Thyristor-based STATCOM” International Conference On Power System Technology 2006
- [9] Chong Han, Member, IEEE, Zhanoning Yang, Bin Chen, Alex Q. Huang, Fellow, IEEE, Bin Zhang, Student Member, IEEE, Michael R. Ingram, Senior Member, IEEE, and Abdel-Aty Edris, Senior Member, IEEE “Evaluation of Cascade-Multilevel-Converter-Based STATCOM for Arc Furnace Flicker Mitigation” IEEE Transactions On Industry Applications, Vol. 43, No. 2, March/April 2007
- [10] Kuang Li, Jinjun Liu, Zhaoan Wang, and Biao Wei “Strategies and Operating Point Optimization of STATCOM Control for Voltage Unbalance Mitigation in Three-Phase Three-Wire Systems” IEEE Transactions On Power Delivery, Vol. 22, No. 1, January 2007
- [11] C. K. Lee, Joseph S. K. Leung, Member, IEEE, S. Y. Ron Hui, Fellow, IEEE, and Henry Shu-Hung Chung, Member, IEEE “Circuit-Level Comparison of STATCOM Technologies” IEEE Transactions On Power Electronics, Vol. 18, No. 4, July 2003
- [12] Diego Soto, Member, IEEE, and Rubén Peña, Member, IEEE “Nonlinear Control Strategies for Cascaded Multilevel STATCOMs” IEEE Transactions On Power Delivery, Vol. 19, No. 4, October 2004
- [13] Y. Chen, B. Mwinyiwiwa, Z. Wolanski, and B.-T. Ooi, “Regulating and qualizing dc capacitance voltages in multi-level STATCOM,” IEEE Trans. Power Del., vol. 12, no. 2, pp. 901–907, Apr. 1997.
- [14] Y. H. Liu, J. Arrillaga, and N. R. Watson, “Multi-level voltage reinjection—A new concept in high power voltage source conversion,” Proc. Gen. Transm. Distrib., vol. 151, no. 3, pp. 290–298, 2004.
- [15] Y. H. Liu, J. Arrillaga, and N. R. Watson, “A new STATCOM configuration using multi-level dc voltage reinjection for high power applications,” IEEE Trans. Power Del., vol. 19, no. 4, pp. 1828–1834, Oct. 2004.

- [16] Y. H. Liu, J. Arrillaga, and N. R. Watson, "Cascaded H-bridge voltage reinjection-Part II: Application to HVDC transmission," Companion Paper.
- [17] Y. H. Liu, J. Arrillaga, and N. R. Watson, "STATCOM performance of a multi-level voltage reinjection converter," presented at the IEEE/Power Eng. Soc. Transmission Distribution Conf., Dalian, China, 2005.

Authors

M. Vishnu Prasad was born in Andhra Pradesh, India, received the B.Tech Electrical and Electronics Engineering from Dr. Paul Raj Engineering college affiliated to JNT University, Hyderabad in the year 2007 and M.Tech .Power Electronics & Drives from SRM University, India in the year 2010. Currently, he is interested to research topics include Power Electronics especially in multi level inverters. He is currently as a Lecturer of Electrical Electronics Engineering Department at Sri Vasavi Institute of Engg & Technology, Nandamuru, Pedana Mandal, Krishna (Dt) Affiliated to JNT University, Kakinada, Andhra Pradesh, India



K. Surya Suresh was born in Andhra Pradesh, India, received the B.Tech Electrical and Electronics Engineering from Sri Sarathi institute of Engg & Technology affiliated to JNT University, Hyderabad and M.Tech .Power Electronics as concentration from KL University, India. Currently, he is interested to research topics include Power Electronics, multi level inverters and fuzzy logic controllers. He is currently as a Lecturer of Electrical Electronics Engineering Department at Sri Vasavi Institute of Engg & Technology, Nandamuru, Pedana Mandal, Krishna (Dt) Affiliated to JNT University, Kakinada, Andhra Pradesh, India



IMAGE RETRIEVAL USING TEXTURE FEATURES EXTRACTED USING LBG, KPE, KFCG, KMCG, KEVR WITH ASSORTED COLOR SPACES

H.B.Kekre¹, Sudeep D. Thepade², Tanuja K. Sarode³, Shrikant P. Sanas⁴

¹Sr. Professor, ²Associate Professor, ³Associate Professor, ⁴Lecturer ,

MPSTME, SVKM's NMIMS University, Vile Parle (W), Mumbai, India

³Thadomal Shahani Engg. College, Mumbai, India, ⁴Lecturer, RAIT, Mumbai, India

ABSTRACT

This paper presents novel texture based content based image retrieval (CBIR) methods using six assorted color spaces i.e RGB, LUV, YCgCb, YIQ, YCbCr, YUV. For extracting texture feature Vector Quantization (VQ) algorithms like Linde Buzo Gray (LBG), Kekre's Proportionate Error (KPE), Kekre's Error Vector Rotation (KEVR), Kekre's Median Codebook Generation (KMCG) and Kekre's Fast Codebook Generation (KFCG) algorithms are used. The codebook generated from these algorithms act as the feature vector for CBIR system. Seven different codebook sizes varying from 8x12 to 512x12 are generated in six color spaces and using five different VQ algorithms, so in all 210 (7x6x5) techniques are presented here. The proposed image retrieval techniques are tested on generic image database having 1000 images and is independent of the image size and rotation invariant. To test the performance of the proposed CBIR techniques Precision and Recall is used. From the results it is observed that KFCG gives best results with higher precision and recall values. Codebook size 128x12 gives best result in all codebook generation algorithms.

KEYWORDS: CBIR, Vector Quantization, LBG, KPE, KFCG, KEVR, KMCG

I. INTRODUCTION

Visual communication plays an important role in modern communication system. We live in the digital era where the advancement in information and communication technology takes place every day. Large amount of digital data is generated, transmitted, stored, analyzed and accessed. Mostly information is in the form of multimedia nature such as digital images, audio, video and graphics. From various sources large amount of images are generated and it takes large volume to store. This stored information is in the form of images. It is more complex to retrieve and difficult to store in large volume. The need for efficient retrieval of images has been recognized by managers of large image collections. To develop efficient indexing techniques for the retrieval of enormous volumes of images being generated these days, we need to achieve reasonable solutions to these above mentioned problems needs to be achieved. Content based image retrieval gives solution of above problem.

The term 'content' in CBIR refers to colors, shapes, textures, or any other information that can be possibly obtained from the image itself and 'Content Based' denotes that the search will consider the concrete contents of the image. It gives query as Image and output is number of matching images to query image. In a CBIR, features are used to represent the image content. The features are extracted automatically and there is no manual intervention, thus eliminating the dependency on humans in the feature extraction stage. The typical CBIR system performs two major tasks. The first one is feature extraction (FE), where a set of features, called feature vector, is generated to accurately represent the content of each image in the database. A feature vector is much smaller in size than the original image. The second task is similarity measurement (SM), where a distance between the query image

and each image in the database using their signatures is computed so that the top “closest” images can be retrieved. Many current CBIR system use Euclidean distance on the extracted feature set as a similarity measure. The Direct Euclidian distance between image P and query image Q is given in equation 1. where V_{pi} and V_{qi} are the feature vectors of image P and query image Q respectively with size ‘n’.

$$ED = \frac{1}{2} \sqrt{\sum_{i=1}^n (V_{pi} - V_{qi})^2} \quad (1)$$

A variety of feature extraction techniques have been developed. Color based feature extraction techniques include color histogram, color coherence vector, color moments,, circular ring histogram [4,16], BTC extensions [17,18,20]. Texture based feature extraction techniques such as co-occurrence matrix [6], Fractals [2], Gabor filters [2], variations of wavelet transform [1], Kekre transform [7,9,12] have been widely used. Effort has been made in even to extend image retrieval methodologies using combination of color and texture as the case in [8] where Walshlet Pyramids are introduced. The synergy resulting from the combination of color and texture is demonstrated to be superior than using just color and texture [13, 14, 15].

II. TEXTURE FEATURE EXTRACTION METHODS

Texture is important component of human visual perception and can be effectively used for identifying different image regions [1]. Compared with color and shape features, texture features[5] indicate the shape distribution, better suits the macrostructure and microstructure of the images [2]. Texture representation methods can be classified into three categories, namely structural, statistical and multi-resolution filtering methods. The identification of specific textures in an image is achieved primarily by modeling texture as a two-dimensional gray level variation [6, 13]. This two dimensional array is called as Gray level Co-occurrence Matrix (GLCM). GLCM describes the frequency of one gray tone appearing in a specified spatial linear relationship with another gray tone, within the area under investigation.

III. VECTOR QUANTIZATION

Vector Quantization is an efficient technique for data compression [10]. VQ has been very popular in variety of research fields such as video-based event detection, speech data compression, image segmentation, CBIR, face recognition, iris recognition, data hiding etc. VQ can be defined as the mapping function that maps k-dimensional vector space to the finite set $CB = \{ C_1, C_2, C_3, \dots, C_N \}$. The set CB is called codebook consisting of N number of code vectors and each code vector $C_i = \{ ci_1, ci_2, ci_3, \dots, ci_k \}$ is of dimension k. The codebook is the feature vector of the entire image and can be generated by using clustering techniques. Here we calculate results on different size of Codebooks like 8, 16, 32, 64, 128, 256 and 512 are used.

Five assorted algorithms are used in Vector Quantization to generate codebooks alias Linde-Buzo-Gray Algorithm(LBG)[3], Kekre’s Proportionate Error Algorithm (KPE)[17], Kekre’s Error Vector Rotation Algorithm(KEVR)[23], Kekre’s Fast CodeBook Genearion Algorithm(KFCG)[11][19][21], and Kekre’s Median CodeBook Genearion Algorithm (KMCG)[24].

IV. COLOR SPACES

Many attempts have been made to model color perception by researchers of various fields: psychology, perception, computer vision, image retrieval, and graphics. Some of these resulted in well defined color spaces. The list of color spaces is almost endless. A few of the most important color spaces are: RGB (Red, Green, and Blue), HSV (Hue, Saturation, and Value), HIS (also named HSB), rgb, XYZ, Kekre’s YCgCb, Kekre’s LUV, YUV, YIQ, YCbCr. Color spaces are needed in the representation of color-ranges. The manipulation of colors, as is done in the graphics industry. Mixing of colors. The retrieval (i.e. matching) of colors. So it is evident that CBIR engines [22], using color as feature, need a color space for color matching. However, often CBIR engines also use a color space in their color selection interface.

V. RESULTS AND DISCUSSIONS

The CBIR techniques are tested on the augmented Wang image database of 1000 variable size images spread across 11 categories of human beings, animals, natural scenery and man-made things. Image database contains various size of images. Figure 1 shows the sample database of 16 images randomly selected from each category. The images are of varying sizes ranging from 384x256 to 84x128. From each category randomly selected 5 images are treated as queries which are fired for obtaining the results of proposed CBIR methods.

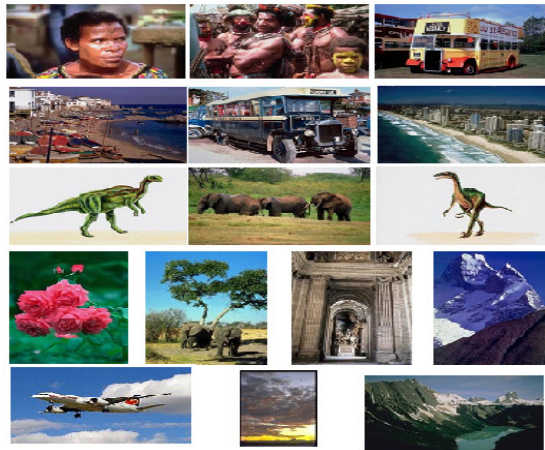


Figure 1: Sample Image from the considered Image Database

The performance of proposed CBIR methods is evaluated using the crossover point value of average precision and average recall values for considered queries. The standard definitions of precision and recall are given by equation 2 and equation 3. Precision gives the accuracy whereas the Recall gives completeness.

$$\text{Precision} = \frac{\text{Number_of_relevant_images_retrieved}}{\text{Total_number_of_images_retrieved}} \quad (2)$$

$$\text{Recall} = \frac{\text{Number_of_relevant_images_retrieved}}{\text{Total_number_of_relevant_images_in_database}} \quad (3)$$

Figure 2 shows the crossover point, the point where precision and recall crosses. This crossover point is used as the performance measure of CBIR techniques. Higher value of the crossover point of precision and recall reflects better performance.

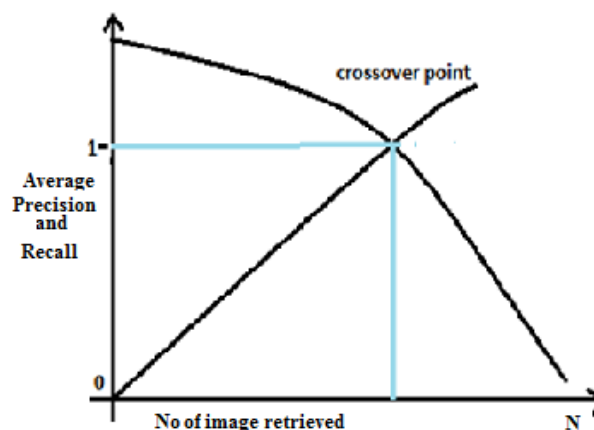


Figure 2. Crossover point of average precision and recall.

Figure 3 gives crossover points of average precision and average recall values of the LBG based CBIR techniques for all considered codebook sizes and different color spaces tested on Wang image database. Here the codebook sizes 16, 32 and 64 are better with higher crossover point values in all color spaces except the RGB color space. The precision and recall crossover point of codebook size 32 for LUV color space is higher than the same for other codebook sizes and color spaces indicating better performance for LBG-CBIR on Wang image database.

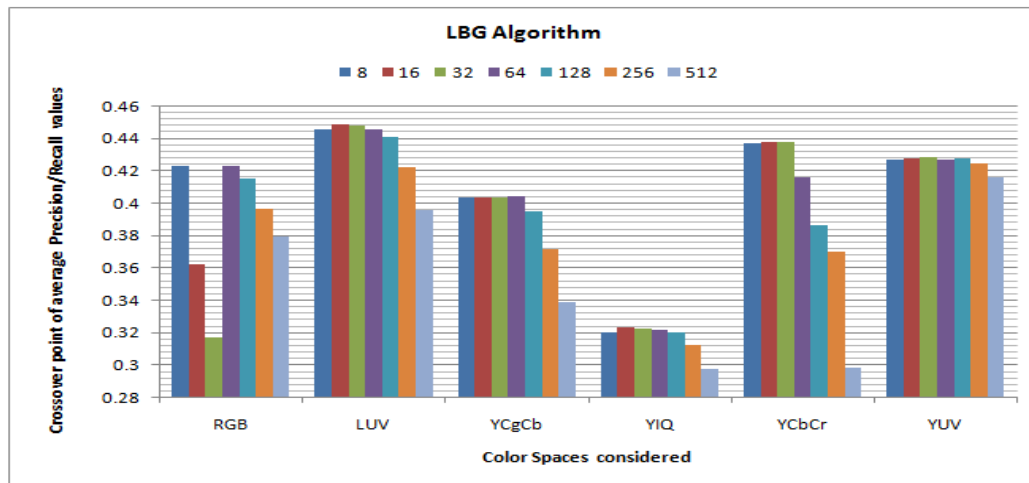


Figure 3: Performance comparison of various color spaces of different Codebook sizes for LBG algorithm using Precision-Recall crossover points.

Figure 4 gives crossover points of average precision and average recall values of the KPE based CBIR techniques for all codebook sizes and different color spaces tested on Wang image database. Here the codebook sizes 8 and 16 are better with higher crossover point values in all color spaces. The precision and recall crossover value of codebook size 8 for YCbCr color space is higher than the same for other codebook sizes and color spaces indicating better performance for KPE-CBIR on Wang image database.

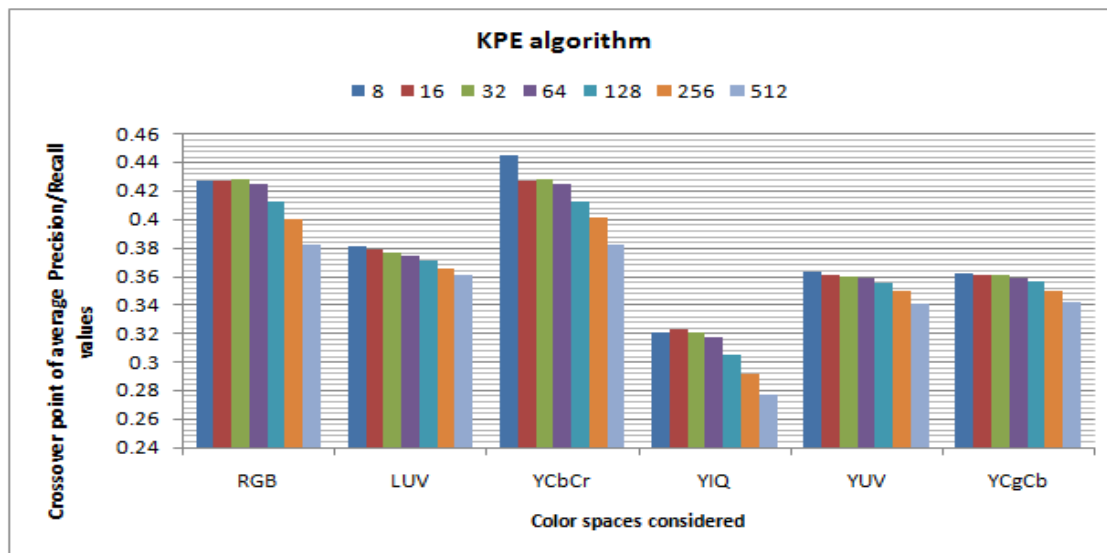


Figure 4 : Performance comparison of various color spaces of different Codebook sizes for KPE algorithm using Precision-Recall crossover points

Figure 5 gives crossover points of average precision and average recall values of the KEVR based CBIR techniques for all codebook sizes and different color spaces tested on Wang image database. Here the codebook sizes 64 and 128 are better with higher crossover point values. The precision and recall crossover point value of codebook size 128 for YCbCr color space is higher than the same for

other codebook sizes and color spaces indicating better performance for KEVR-CBIR on Wang image database.

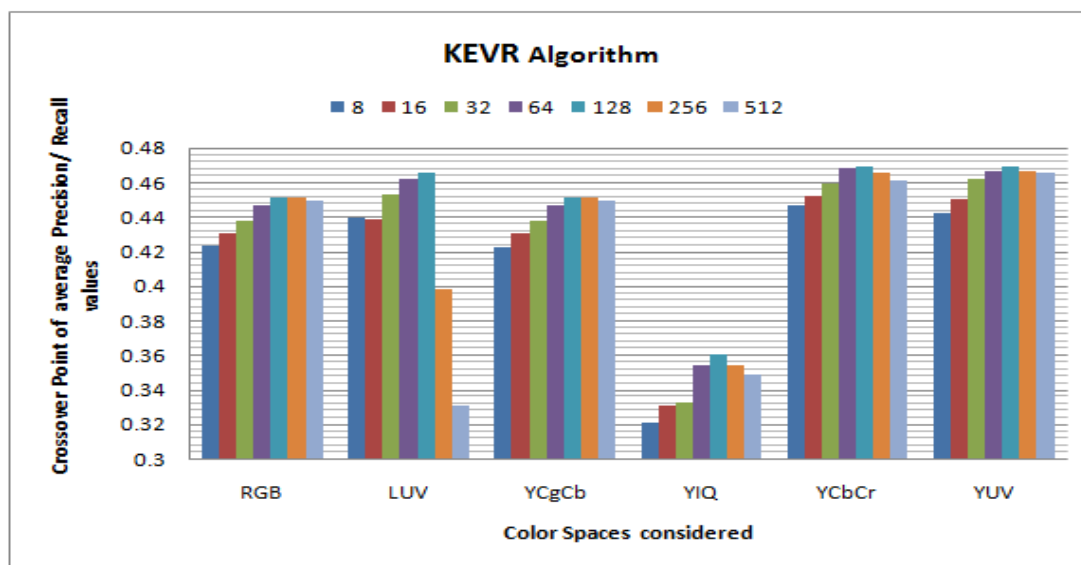


Figure 5 : Performance comparison of various color spaces of different Codebook sizes for KEVR algorithm using Precision-Recall crossover points

Figure 6 gives crossover points of average precision and average recall values of the KFCG based CBIR techniques for all codebook sizes and different color spaces tested on Wang image database. Here the codebook sizes 128 and 256 are better with highest crossover point value. The precision and recall curves of codebook size 256 for RGB color space are higher than other codebook sizes and color spaces indicating better performance for KFCG-CBIR on Wang image database.

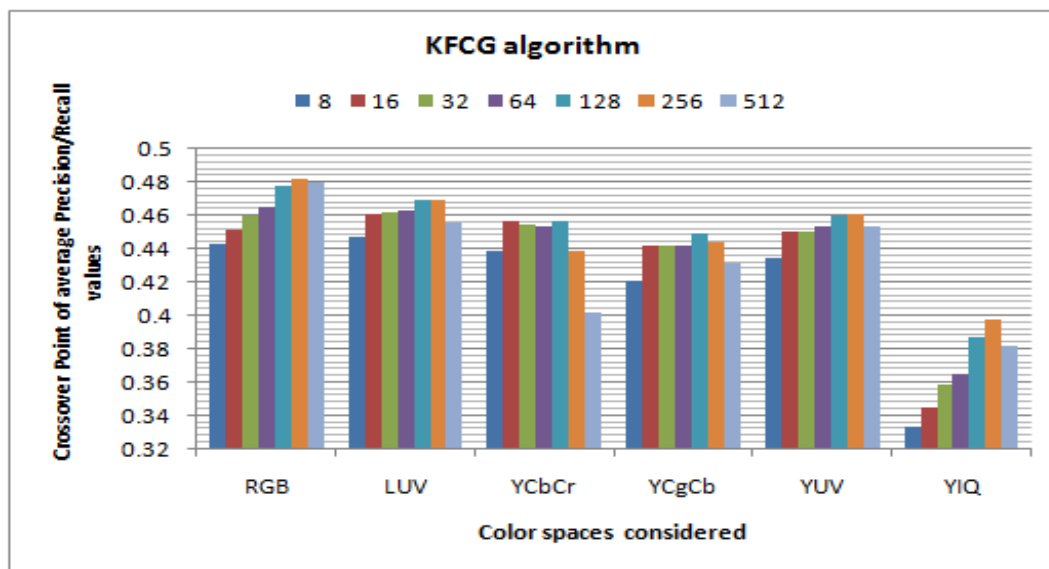


Figure 6 : Performance comparison of various color spaces of different Codebook sizes for KFCG algorithm using Precision-Recall crossover points

Figure 7 gives crossover points of average precision and average recall values of the KMCG based CBIR techniques for all codebook sizes and different color spaces tested on Wang image database. Here the codebook sizes 128 and 256 are better with highest crossover point value. The precision and recall curves of codebook size 128 for LUV color space are higher than other codebook sizes and color spaces indicating better performance for KMCG-CBIR on Wang image database.

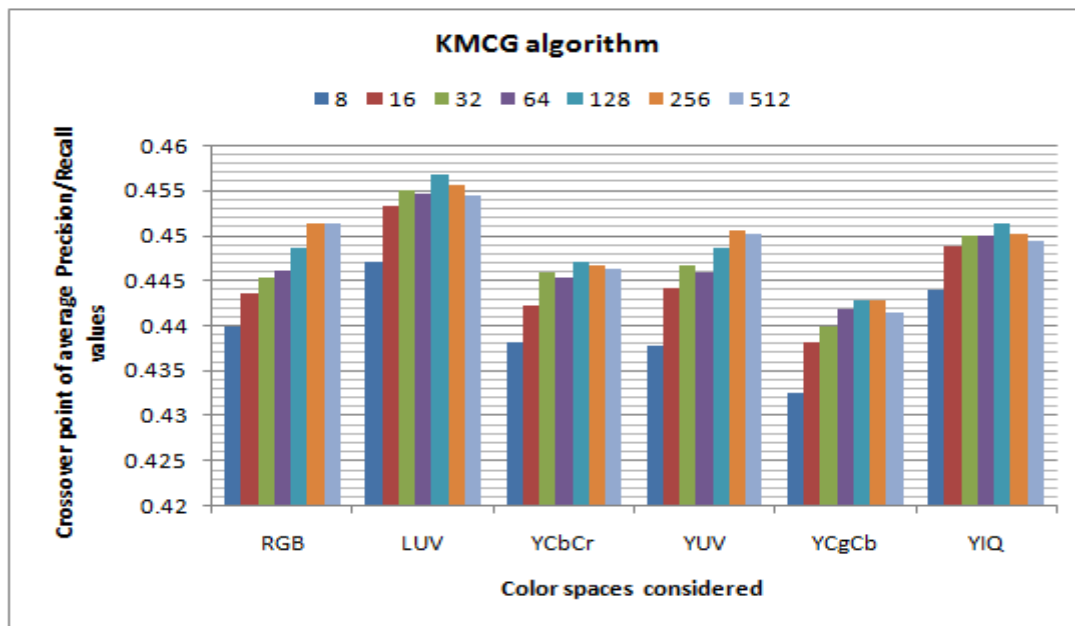


Figure 7: Performance comparison of various color spaces of different Codebook sizes for KMCG algorithm using Precision-Recall crossover points

5.1: Results according to Different Color Spaces.

Her proposed CBIR methods with respect to various color spaces are discussed for respective codebook generation techniques.

Figure 8 shows the comparison between various sizes of codebook for RGB color spaces. Codebook size 128 and 256 gives the better result than other size of codebook. After codebook size 256 graphs is slightly decreasing because more no of voids are created. KFCG gives the better result as compared to other VQ methods.

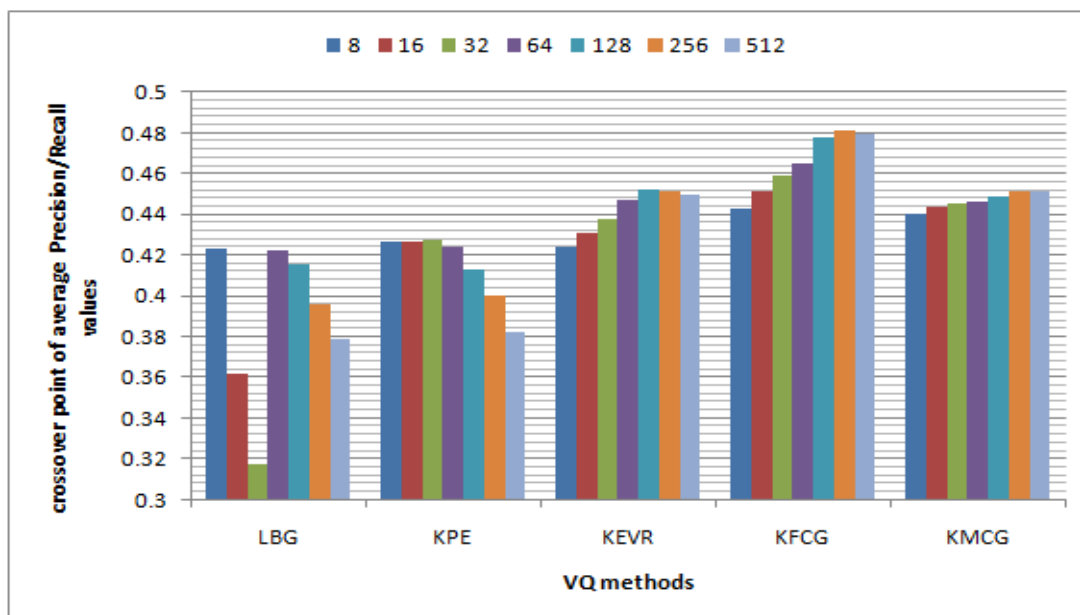


Figure 8 : Crossover points of Avg Precision & Recall plotted against the various VQ methods for RGB color space.

Figure 9 shows the comparison between various sizes of codebook for LUV color spaces. Codebook size 128 gives the better result than other size of codebook. KFCG gives the better result as compared to other VQ methods.

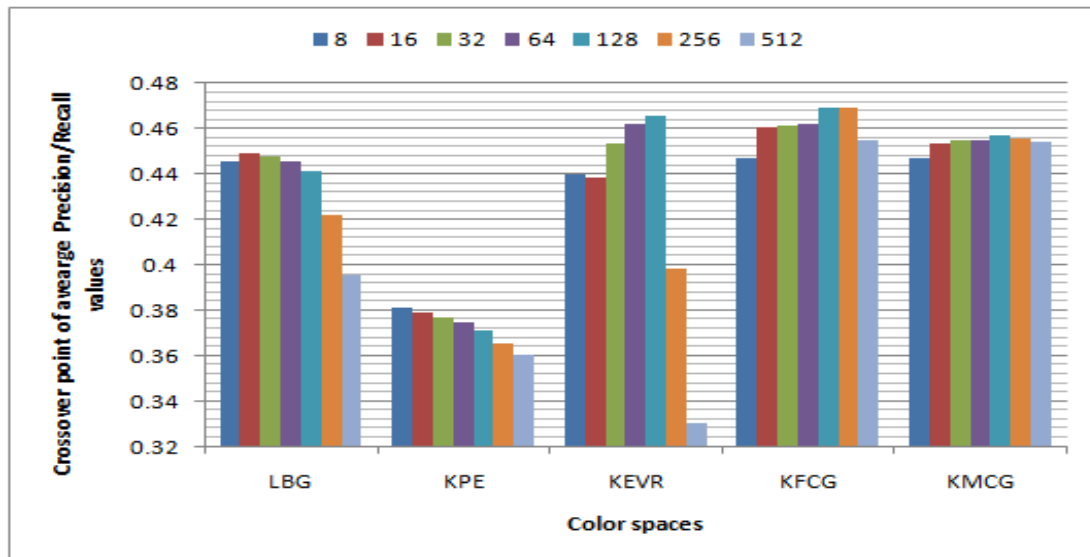


Figure 9 : Crossover points of Avg Precision & Recall plotted against the various VQ methods for LUV color space.

Figure 10 shows the comparison between various sizes of codebook for YUV color spaces. Codebook size 128 and 256 gives the better result than other size of codebook. KEVR gives the better result as compared to other VQ methods.

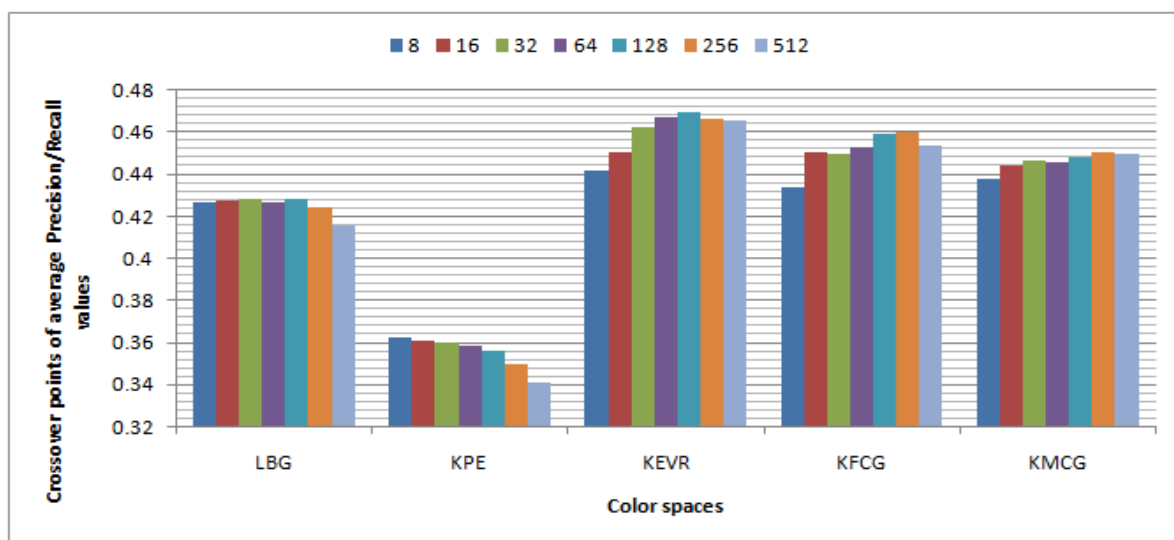


Figure 10: Crossover points of Avg Precision & Recall plotted against the various VQ methods for YUV color space.

Figure 11 shows the comparison between various sizes of codebook for YCbCr color spaces. Codebook size 64 and 128 gives the better result than other size of codebook. KEVR gives the better result as compared to other VQ methods.

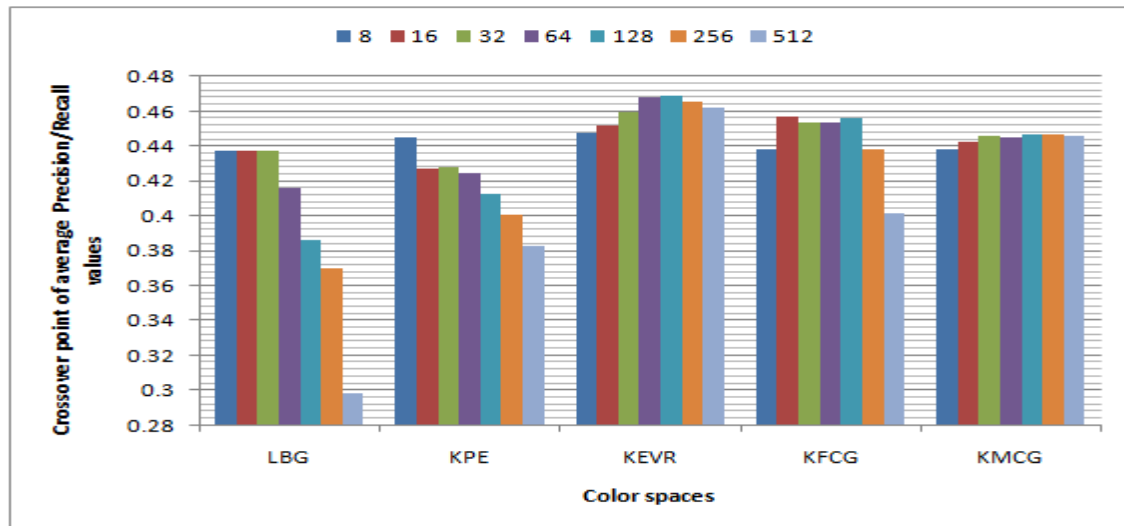


Figure 11: Crossover points of Avg Precision & Recall plotted against the various VQ methods for YCbCr color space.

Figure 12 shows the comparison between various sizes of codebook for YCgCb color spaces. Codebook size 128 and 256 gives the better result than other size of codebook. KEVR gives the better result as compared to other VQ methods.

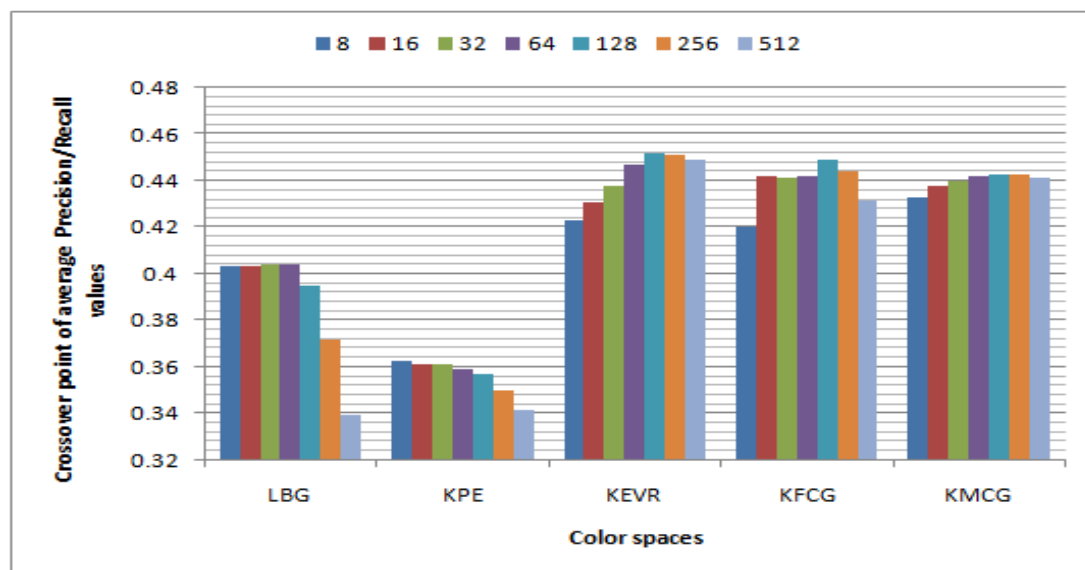


Figure 12: Crossover points of Avg Precision & Recall plotted against the various VQ methods for YCgCb color space.

Figure 13 shows the comparison between various sizes of codebook for YIQ color spaces. Codebook size 128 gives the better result than other size of codebook. KMCG gives the better result as compared to other VQ methods.

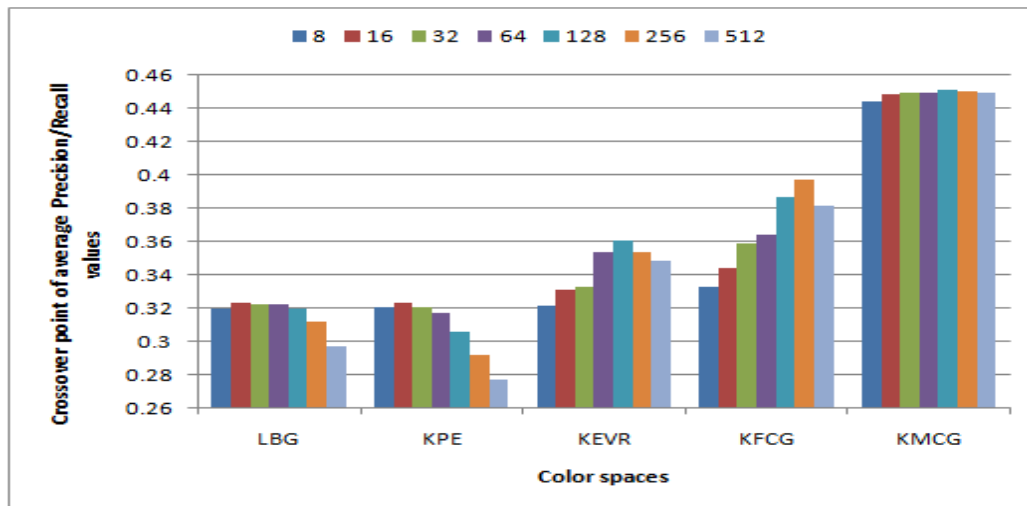


Figure 13: Crossover points of Avg Precision & Recall plotted against the various VQ methods for YIQ color space.

5.2. Discussion.

Table 1 shows the best Crossover values for respective color spaces and codebook sizes with respect to VQ algorithms LBG, KEVR, KPE, KFCG, KMCG. It is observed that LBG algorithm in YUV space gives higher crossover point for codebook size 16. In LBG-CBIR YUV color space gives better result in all size of codebook. YIQ gives worst result compare to other color spaces. KPE algorithm in YCbCr color space gives better result for codebook size 8. In KPE-CBIR YCbCr color space gives average result and YIQ color space gives worst result as compare to other color spaces. KEVR algorithm in YUV color space gives better results for codebook size 128. In KEVR-CBIR LUV and YCbCr color space gives average result and YIQ color space gives worst result as compare to other color spaces. KMCG algorithm in LUV color space gives better result for codebook size 128. In KFCG algorithm RGB color space gives better result for codebook size 256 where as LUV color space gives average result while YIQ and YUV color space gives worst results.

Table 1. Best Crossover values for respective color space and codebook size with respect to various VQ algorithms.

Algorithm	Color space	codebook size	crossover point
LBG	LUV	16	0.44885
KEVR	YUV	128	0.469673
KPE	YCbCr	8	0.4445
KMCG	LUV	128	0.45675
KFCG	RGB	256	0.4815

Table 2 shows the best Crossover values for various algorithms and codebook sizes with respect to various color spaces. It is observed from the table that KFCG algorithm gives highest crossover point in RGB color space for codebook size 256. KEVR and KMCG gives average result where as LBG gives worst results in RGB color space as compare to other algorithms. LUV color space gives better results for codebook size 128 of KFCG. KMCG and KEVR algorithm gives average result and KPE algorithm gives worst results in LUV space. In YUV color space KEVR algorithm gives higher crossover point for codebook size 128 where as KFCG and KMCG gives average results and KPE algorithm gives worst result. KEVR algorithm gives better result for codebook size 128 in YCbCr color space where as KFCG and KMCG algorithm give average result while LBG algorithm gives worst result as compare to others algorithm. In YCgCb color space KEVR algorithm gives better result for codebook size 128. KFCG and KMCG algorithm give average result while KPE algorithm gives worst result in YCgCb color space. KMCG algorithm gives better result for codebook size 128 in YIQ color space while KFCG gives average result where as KPE gives worst result. As compare to

all the results RGB color space Kerkre's fast codebook generation algorithm gives best performance crossover value is 0.4814 for codebook size 256.

Table 2 : Best Crossover values for various algorithms and codebook sizes with respect to various color spaces.

Color space	Algorithm	codebook size	crossover point
RGB	KFCG	256	0.4815
LUV	KFCG	128	0.46905
YUV	KEVR	128	0.46967
YCbCr	KEVR	128	0.46916
YCgCb	KEVR	128	0.4517
YIQ	KMCG	128	0.45125

VI. CONCLUSION

The performance of CBIR system depends on the precision and recall. The crossover point of precision and recall is taken as criteria for judging the performance of CBIR technique. The use of vector quantization codebooks as feature vectors for image retrieval is proposed in the paper. The codebook generation techniques such as of Linde-Buzo-Gray (LBG), Kerkre's Proportionate Error Algorithm(KPE), Kerkre's Fast Codebook Generation algorithm, Kerkre's Median Codebook Generation algorithm and newly introduced Kerkre's Error Vector Rotation (KEVR) algorithms for texture feature extraction are used. These codebooks extracted with sizes 8, 16, 32, 64, 128, 256 and 512 are used in proposed CBIR techniques. Six different color spaces are used. Thus the five codebook generation algorithms, seven different codebook sizes and six color spaces per algorithm results into 210 variations of proposed image retrieval techniques. All these variations are tested on Wang image database of 1000 images. As compared to all the discussed CBIR variations results Kerkre's fast codebook generation algorithm in RGB color space gives best performance crossover value is 0.4814 for codebook size 256.

REFERENCES

- [1] Sanjoy Kumar Saha, Amit Kumar Das, Bhabatosh Chanda, "CBIR using Perception based Texture and Color Measures", in Proc. of 17th International Conference on Pattern Recognition(ICPR'04), Vol. 2, Aug 2004.
- [2] Xiaoyi Song, Yongjie Li, Wufan Chen,"A Textural Feature Based Image Retrieval Algorithm", in Proc. of 4th Int. Conf. on Natural Computation, Oct. 2008.
- [3] H.B.Kekre, Sudeep D. Thepade, Tanuja K. Sarode and Vashali Suryawanshi 'Image Retrieval using Texture Features extracted from GLCM, LBG and KPE'. International Journal of Computer Theory and Engineering, Vol. 2, No. 5, October, 2010
- [4] Wang Xiaoling, "A Novel Circular Ring Histogram for Content-based Image Retrieval", First International Workshop on Education Technology and Computer Science, 2009.
- [5] Xiaoyi Song, Yongjie Li, Wufan Chen,"A Textural Feature Based Image Retrieval Algorithm", in Proc. of 4th Int. Conf. on Natural Computation, Oct. 2008.
- [6] Jing Zhang, Gui-li Li, Seok-wum He, "Texture-Based Image Retrieval By Edge Detection Matching GLCM ", in Proc. of 10th International conference on High Performance Computing and Comm., Sept. 2008.
- [7] H.B.Kekre, Sudeep D. Thepade, "Panoramic View Construction using Partial Images", IEEE sponsored International Conference on Sensors, Signal Processing, Communication, Control and Instrumentation (SSPCCIN-2008) 03-05 Jan 2008, Vishwakarma Institute of Technology, Pune.
- [8] H. B. Kekre, Tanuja K. Sarode, "Speech Data Compression using Vector Quantization", WASET International Journal of Computer and Information Science and Engineering (IJCISE), vol. 2, No. 4, pp.: 251-254, Fall 2008. available: <http://www.waset.org/ijcise>.
- [9] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Non-Involutorial Orthogonal Kerkre's Transform", International Journal of Multidisciplinary Research and Advances in Engineering (IJMRAE), Ascent Publication House, 2009, Volume 1, No.I, 2009. Abstract available online at www.ascent-journals.com.

- [10] Khalid Sayood ,” Introduction to Data Compression ,” University of Nebraska-Lincoln , Second Edition , ISBN:1-55860-558-4, by Academic Press, 2000.
- [11] H.B.Kekre, Archana Athawale, Tanuja sarode, Mixing Codebooks of LBG, KPE and KFCG Algorithms to Increase Capacity of Information Hiding international Journal of Computer Applications (0975 – 8887) Volume 5– No.3, August 2010.
- [12] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke, “Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors”, International Journal of Computer Science and Network Security (IJCSNS), Volume:10, Number 1, January 2010, (ISSN: 1738-7906) Available at www.IJCSNS.org.
- [13] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, Vaishali Suryavanshi, “Improved Texture Feature Based Image Retrieval using Kekre’s Fast Codebook Generation Algorithm”, Springer-International Conference on Contours of Computing Technology (Thinkquest-2010), Babasaheb Gawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.
- [14] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, “Image Retrieval by Kekre’s Transform Applied on Each Row of Walsh Transformed VQ Codebook”, (Invited), ACM-International Conference and Workshop on Emerging Trends in Technology (ICWET 2010), Thakur College of Engg. And Tech., Mumbai, 26-27 Feb 2010, The paper is invited at ICWET 2010. Also will be uploaded on online ACM Portal.
- [15] H.B.Kekre, Tanuja Sarode, Sudeep D. Thepade, “Color-Texture Feature based Image Retrieval using DCT applied on Kekre’s Median Codebook”, International Journal on Imaging (IJI), Volume 2, Number A09, Autumn 2009, pp. 55-65. Available online at www.ceser.res.in/iji.html (ISSN: 0974-0627).
- [16] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, “Query by image and video content: the QBIC system”. IEEE Computer, 28(9):23–32, 1995.
- [17] H. B. Kekre, Kamal Shah, Tanuja K. Sarode, Sudeep D. Thepade, ”Performance Comparison of Vector Quantization Technique – KFCG with LBG, Existing Transforms and PCA for Face Recognition”, International Journal of Information Retrieval (IJIR), Vol. 02, Issue 1, pp.: 64-71, 2009.
- [18] H.B.Kekre, Sudeep D. Thepade, “Color Based Image Retrieval using Amendment Block Truncation Coding with YCbCr Color Space”, International Journal on Imaging (IJI), Volume 2, Number A09, Autumn 2009, pp. 2-14. Available online at www.ceser.res.in/iji.html (ISSN: 0974-0627).
- [19] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, Vaishali S., “Improved Texture Feature Based Image Retrieval using Kekre’s Fast Codebook Generation Algorithm.” In: Springer-Int. Conf. on Contours of Computing Tech. (Thinkquest-2010), 13-14 March, BGIT, Mumbai (2010).
- [20] H.B.Kekre, Sudeep D. Thepade, “Using YUV Color Space to Hoist the Performance of Block Truncation Coding for Image Retrieval”, In Proc. of IEEE International Advanced Computing Conference 2009 (IACC’09), Thapar University, Patiala, INDIA, 6-7 March 2009.
- [21] H. B. Kekre, Tanuja K. Sarode, Sudeep D. Thepade, “Image Retrieval using Color-Texture Features from DCT on VQ Codevectors obtained by Kekre’s Fast Codebook Generation.” In: ICGST-Int. Journal GVIP, Vol. 9, Issue 5, pp. 1-8, (Sept 2009).
- [22] S. Jain, A. Joglekar, and D. Bote, ISearch: “A Content-based Image Retrieval (CBIR) Engine”, as Bachelor of Computer Engineering final year thesis, Pune University, 2002.
- [23] H. B. Kekre, Tanuja K. Sarode “New Clustering Algorithm for Vector Quantization using Rotation of Error Vector”, International Journal of Computer Science and Information Security, Vol. 7, No. 03, 2010, USA.
- [24] H. B. Kekre, Tanuja K. Sarode, “An Efficient Fast Algorithm to Generate Codebook for Vector Quantization,” First International Conference on Emerging Trends in Engineering and Technology, ICETET-2008, held at Rasoni College of Engineering, Nagpur, India, pp.: 62- 67, 16-18 July 2008. Available at IEEE Xplore.

Author Biographies

H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering. College, Mumbai. Now he is Senior Professor at MPSTME, SVKM’s NMIMS. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 270 papers in National / International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and



Life Member of ISTE Recently seven students working under his guidance have received best paper awards. Currently 10 research scholars are pursuing Ph.D. program under his guidance.

Sudeep D. Thepade has Received B.E.(Computer) degree from North Maharashtra University with Distinction in 2003. M.E. in Computer Engineering from University of Mumbai in 2008 with Distinction, Ph.D. from SVKM's NMIMS in 2011, Mumbai. He has about 09 years of experience in teaching and industry. He was Lecturer in Dept. of Information Technology at Thadomal Shahani Engineering College, Bandra(w), Mumbai for nearly 04 years. Currently working as Associate Professor and HoD Computer Engineering at Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS, Vile Parle(w), Mumbai, INDIA. He is member of International Advisory Committee for many International Conferences, acting as reviewer for many referred international journals/transactions including IEEE and IET. His areas of interest are Image Processing and Biometric Identification. He has guided five M.Tech. projects and several B.Tech projects. He more than 125 papers in National/International Conferences/Journals to his credit with a Best Paper Award at International Conference SSPCCIN-2008, Second Best Paper Award at ThinkQuest-2009, Second Best Research Project Award at Manshodhan 2010, Best Paper Award for paper published in June 2011 issue of International Journal IJCSIS (USA), Editor's Choice Awards for papers published in International Journal IJCA (USA) in 2010 and 2011..



Tanuja K. Sarode has Received Bsc.(Mathematics) from Mumbai University in 1996, Bsc.Tech.(Computer Technology) from Mumbai University in 1999, M.E. (Computer Engineering) degree from Mumbai University in 2004, Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 11 years of experience in teaching. Currently working as Assistant Professor in Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is life member of IETE, member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT), Singapore. Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has more than 100 papers in National /International Conferences/journal to her credit.



Shrikant Sanas has received B.E. (Computer) degree from Mumbai University with First Class in 2008. M.-Tech. in Computer Engineering from Mukesh Patel School of Tech. Mgmt. and Engineering, SVKM's NMIMS University Mumbai. in 2011 with Distinction. Currently working as Lecturer in Ramrao Adik Institute of Technology. Nerul, Navi Mumbai. His areas of interest are Image Processing and Database. He has 05 papers in International Journals.



DEVELOPMENT OF A SIMPLE & LOW-COST INSTRUMENTATION SYSTEM FOR REAL TIME VOLCANO MONITORING

Didik R. Santoso, Sukir Maryanto and A.Y. Ponco Wardoyo
Department of Physics, Brawijaya University, Jl. Veteran 2 Malang 65145, Indonesia

ABSTRACT

An instrumentation which is used to monitoring of volcano activities usually has a complex structure and expensive. This, may difficult not only in procurement of the device but also in terms of maintenance. Aim of this research is to develop a simple and inexpensive instrumentation system for online volcano monitoring. For this purpose, the instrumentation system is built upon two main units, i.e. Remote Terminal Unit (RTU) and Master Terminal Unit (MTU). The RTU is sensing unit, which includes seismic sensor module and weather sensor module. Both the sensor modules are equipped with a microcontroller based data acquisition system. The MTU is control and data logger unit. It is built based on PC, and installed application's software for data logger and interface to the internet network, allowing users to access the volcano activity that was monitored by real time, from anywhere. The connection between MTU and RTU performed wirelessly using a digital radio transceiver. The RTU's work function is fully controlled by the MTU. This system has been tested on laboratory scale and work well.

KEYWORDS: *Volcano monitoring, Instrumentation, Seismic Sensor, Weather sensor*

I. INTRODUCTION

In country, that has many volcanoes such as Indonesia (see Figure 1), continuously monitoring of volcano activity becomes very important. There are at least two reasons, first is to monitor the level of volcano hazard, relates to mitigation and management of natural disasters particularly those caused by volcanic eruptions to reduce losses and damages. Secondly is to understand the physical processes that occur inside the volcano, such as magma migration and the mechanisms that exist within the volcano, this is more toward advancing the science of volcanoes (volcanology) itself [1].

There are many physical phenomena resulting from the internal behaviour of volcanoes, including external signs that can be measured with instruments when the magma moves or its chemical composition changes, or when its pressure or temperature varies. Then, many methods have been utilized for observing and monitoring the volcano activity and the efforts for establishing a good and reliable system for monitoring and predicting the volcano eruption is never ended. Instruments will never replace the expertise of the volcanologists in charge of the surveillance, but they definitely can help them in taking swift decisions in case of a crisis such as an eruption [3].

The study of earthquakes is one of the most common methods in order to monitor the volcanic activity, with a great success. The whole idea is based on the theory saying that magma causes small earthquakes during trying to find an exit to the surface [4]. Earthquake activity beneath a volcano almost always increases before an eruption because magma and volcanic gas must first force their way up through shallow underground fractures and passageways.



Figure 1. Major volcanoes in Indonesia [2]

Currently, the development of earthquakes sensor (i.e. seismic) leads to technology of MEMS accelerometer. The technology of very small mechanical device driven by electricity or vice versa, and packaged in Integrated Circuit (IC) chips. Advantages of using MEMS for seismic sensor are small size and compact, sensitive, mild, and relatively cheap. MEMS accelerometer is also available in 3-axis sensing (xyz), so it can be used to measure seismic vibrations in three different directions simultaneously [5,6,7]. Earthquake recorders typically record data at 100 sps and 200 sps, meaning that frequencies above 50Hz or 100Hz are not recorded. More can be learnt about earthquakes by using MEMS accelerometers and recorders capable of sampling at up to 2000 sps [8].

Temperature is one of the physical parameters of a volcano that register characteristic increase during reactivation periods. Thermal monitoring is so one of the most important elements of an integral monitoring system. The monitoring of the surface temperature of volcanoes is one of the essential elements to know the state of their volcanism. The surface temperature distribution of a volcano can be observed by using many kinds of temperature sensor, and also with various methods [9,10,11].

On the other hand, in view of the instrumentation system, the use of a microcontroller to build of the system has been widely applied. Implementation of a microcontroller into the hardware design will be able to increase the capabilities and simplify the system. The advantages of using a microcontroller are small dimensions, programmable, simple, reliable and relatively cheap [12]. A sensor can be interfaced to the others functional devices using a low-cost microcontroller and few resources, as in fieldbus sensors [13]. To realize a web sensor, Internet Protocols must be provided and most of the cost is in full-compliant implementation of these high-level protocols (TCP/IP or HTTP). This approach allows access to the sensor from everywhere with a commercial browser instead of a proprietary interface [14].

Usually, the instrumentation system for volcano monitoring has a complex structure and expensive, this may make difficulties not only in device procurement but also in terms of maintenance. So that, building the system in the simple form and low-cost in budget is a challenge for researchers in the field of volcanology. This paper discusses the design and construction of a simple and low-cost instrumentation system for monitoring seismic activity of the volcano and the surrounding weather by online and real-time based on internet network.

II. PROPOSED INSTRUMENTATION SYSTEM

In general, an instrumentation system consists of four main elements; they are sensor, signal conditioning, signal processing, and display [15]. In this research, block diagram of the proposed instrumentation system can be seen in Figure 2. The system is built upon two main units, namely Remote Terminal Unit (RTU) and Master Terminal Unit (MTU). The RTU is made from several modules, i.e. seismic sensor module, weather sensor module, data acquisition (DAQ) module, and data communication module (i.e. RF-transceiver). While the MTU built based on PC, and installed

application's software for data logger and network interface. Communication between RTU and MTU performed wirelessly by using radio transceiver. The system is built based on internet network, allowing users to access the volcano activity that was monitored by online and real time, from anywhere.

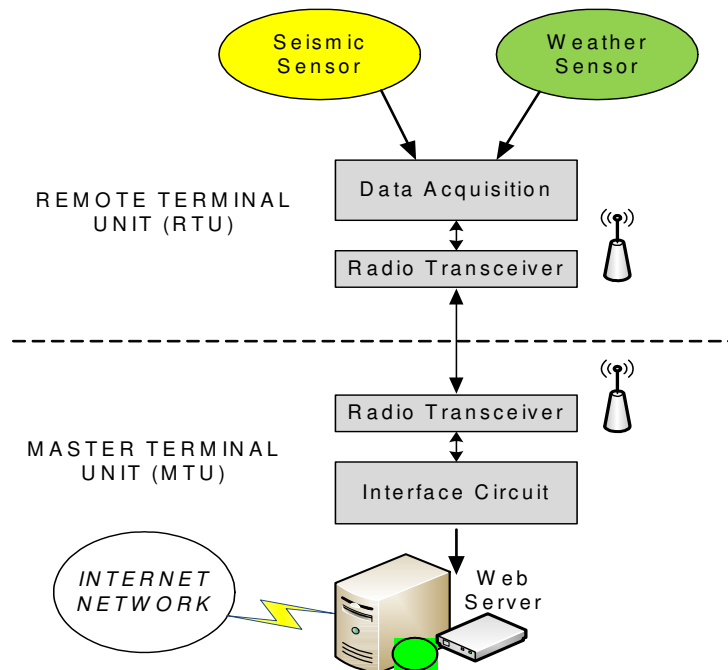


Figure 2. Block diagram of the proposed instrumentation system

2.1 Seismic Sensor

The seismic sensors was developed using MMA7260QT MEMS accelerometer as main component. It is triaxial acceleration sensor, packaged on IC's chip, and produced by Freescale. The MMA7260QT is low-cost capacitive accelerometer, features temperature compensation and g-select which allows for the selection among 4 sensitivities (1.5, 2, 4 and 6 g). Characteristics of this device, in stationary conditions (no motion), output voltage of the sensor is issued a certain volt; depend on position of the sensor placement [16].

For application as seismic sensor, a static voltage is not required and need to be removed. Thus, if the sensor is in static condition, output of the sensor is 0 volt -for all channels (xyz). Then the output of the sensor will correspond to the seismic vibrations only. For the purpose of signal filtering, amplifying, and level adjustment, it is necessary to build a suitable signal conditioning circuit. The signal conditioning consists of a band pass filter (BPF), voltage amplifier and buffer, and voltage level adjustment (DC-offset). BPF circuit serves to eliminate the static DC output voltage of the sensor and blocks high frequency noise. Circuit of the seismic sensors that have been developed given in Figure 3. Transfer function of the circuit is given by:

$$V_{out} = \left(\frac{R_2}{R_1} \right) \left(\frac{C_{in}}{C_f} \right) \left(\frac{1}{(1 + (j\omega/\omega_H)) \left(1 + \frac{1}{(j\omega/\omega_L)} \right)} \right) V_{in} + V_{offset} \quad (1)$$

where,

$$\omega_L = \frac{1}{R_f C_f} \quad \text{then} \quad f_L = \frac{1}{2\pi R_f C_f} \quad \text{Low frequency cut-off} \quad (2)$$

$$\omega_H = \frac{1}{R_{in} C_{in}} \quad \text{then} \quad f_H = \frac{1}{2\pi R_{in} C_{in}} \quad \text{High frequency cut-off} \quad (3)$$

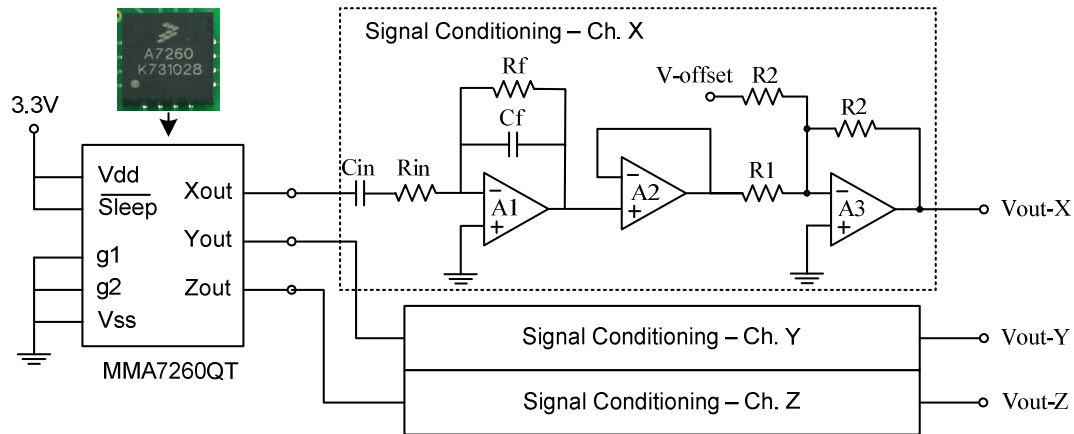


Figure 3. Circuit of 3-axis seismic sensor circuit

Prototype of the sensor in a printed circuit board (PCB) is given in Figure 3. In this figure, left and middle pictures are looked top and bottom sides of the sensor, while the right is a picture of the sensor after casing assembled.

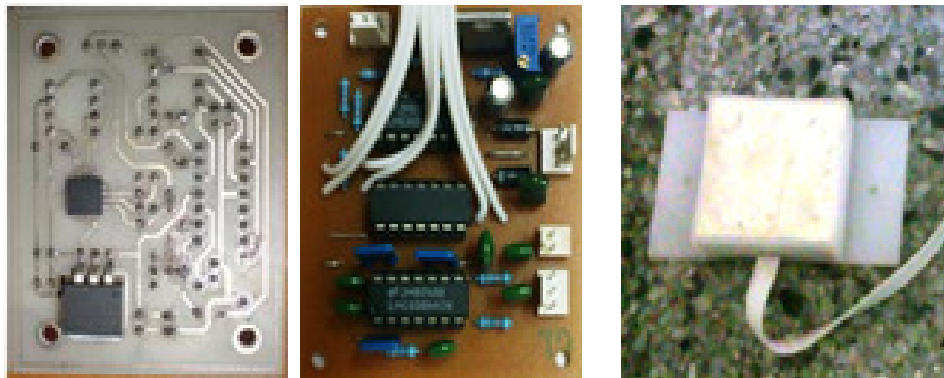


Figure 4. Photos of 3-axis seismic sensor

2.2 Temperatur and Humidity Sensor

Measurement of air temperature and air humidity is important in volcano monitoring. This carried out to determine the environmental changes resulting from the volcano's internal processes. For this purpose, we develop temperature and humidity sensors in simple form by using SHT11. The SHT11 is produced by Senserion, it has double function i.e. as temperature sensor and humidity sensor. The SHT11 has small dimension, high accuracy, and output in digital logic [17]. Another advantage of this sensor is can be directly connected to the I/O port microcontroller without any other additional devices. For temperature functionality, output of the sensors is 14 bits digital data. While for humidity functionality, output of the sensor is 12 bits digital data. Figure 5 shows a picture of this sensor and its connection to the microcontroller.

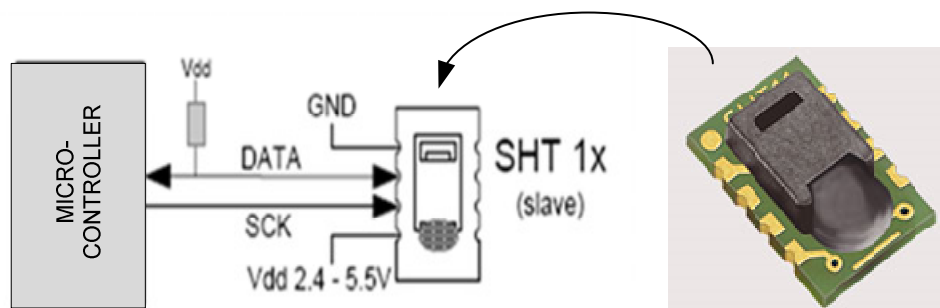


Figure 5. Temperature and humidity sensor

2.3 Wind Direction and Wind Speed Sensor

Wind direction sensor is made using standard mechanical vane system, as shown in Figure 6. Wind direction coded to 4 bits combinational binary code from '0000' to '1111' (Figure 6a), and each bit performed by optocouple circuit (Figure 6b). From these codes, it can be determined of 16 stages different direction, by resolution of 22.5 degrees. Wind speed sensor is created using an anemometer cub system. Magnitude of the wind speed is determined by counting the number of cycles per unit time. The cycle counter performed by optocouple circuit. Prototype of wind direction and wind speed sensor is given in Figure 7.

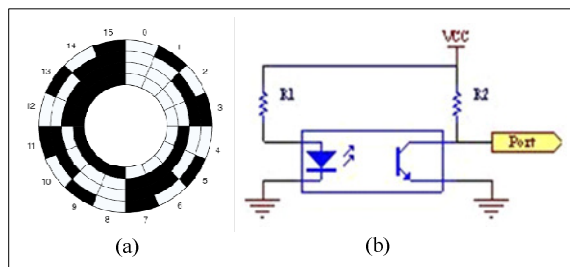


Figure 6. Optocouple circuit

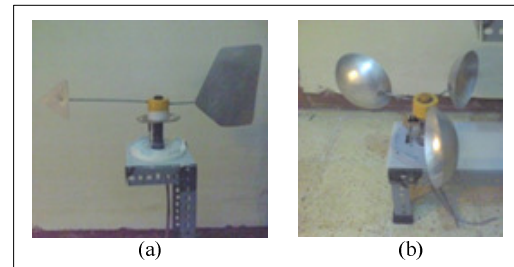


Figure 7. Prototype of (a) wind-direction, (b) wind-speed sensor

2.4 Data Acquisition (DAQ) System

Data acquisition system, built upon hardware and software. The hardware of DAQ system is made by the use of PIC 16F876 microcontroller as main component. PIC 16F876 is midrange microcontroller manufactured by Microchip Company [18]. The advantages of using this device are cheap, has 5-channels internal ADC, and widely available in the commercial market. Figure 8 shows hardware of DAQ system and its connection to the sensors system and the RF transceiver. RF transceiver is radio communication device to perform wireless communication between RTU and MTU. In this research we use YS-320H RF-transceiver from Shenzhen Yishi Electronic Ltd. This is 5 watt (range up to 10 km) wireless data modem [19]. Figure 9 is implementation hardware of (a) DAQ system in PCB and (b) YS-320H RF-transceiver.

Furthermore, the system software of DAQ was made for two purposes, i.e. to hardware system driving and to make display for user interface application. Microcontroller software (i.e. firmware) is constructed by using assembly language MPLAB-IDE, while the display for the user interface was developed using Delphi programming language. Procedure of communication between the RTU (sensing unit) and the MTU (control unit) arranged by program procedure that has been installed both in the microcontroller and PC.

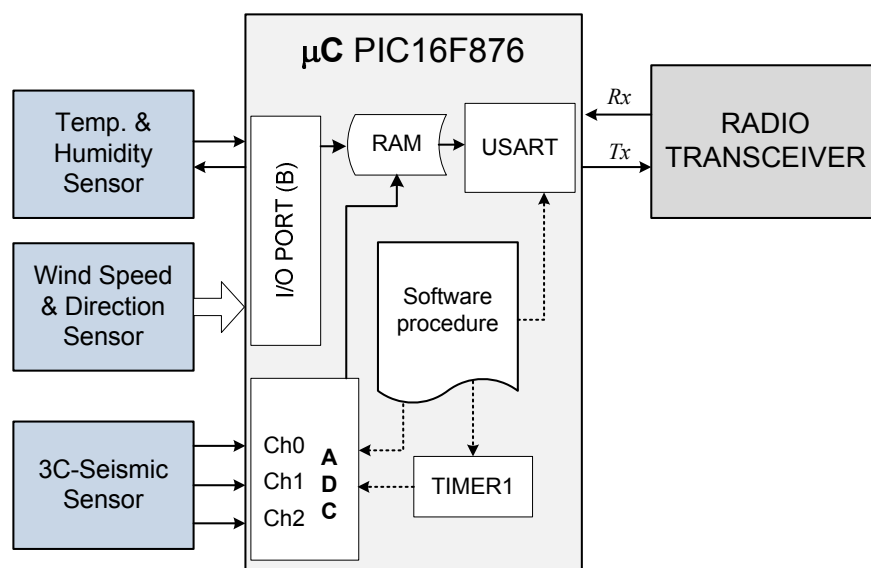


Figure 8. Block diagram of DAQ (in RTU)

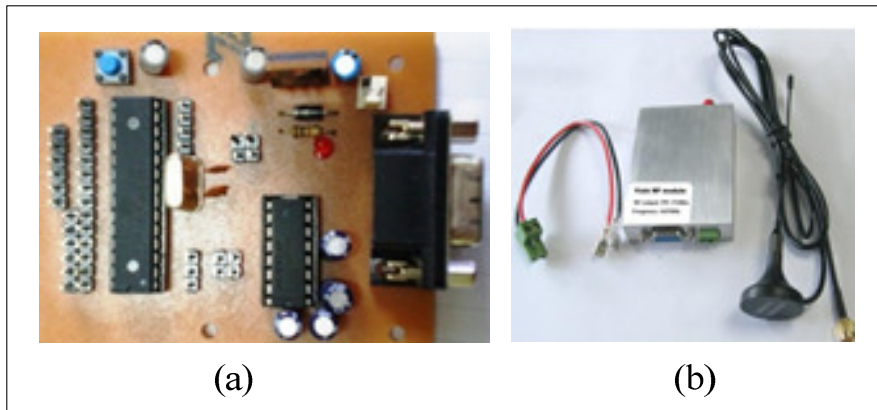


Figure 9. (a) DAQ, (b) RF-transceiver

2.5 Data Logger Device and WEB Interface

Device of data logger is part of the MTU. The data logger serves to store the results of monitoring data come from RTU. Hardware of the data logger is a PC equipped by additional hardware and software interface. In addition, software of the data logger created by using MySQL database application. It's internet based database software to aim that monitoring data can be displayed through the application site, or converted into a spreadsheet.

The acquisition software works by capture the monitoring data in PC's memory that passed on by doing a query insert into the database. The results of acquisitions that have been stored in MySQL database will be processed by components of the application site to real-time concept. Test of data storage prepared to ensure that the data acquisition software can forward the data received by the MTU into the MySQL database table in accordance to sensing time. There is no data queue, data stack (accumulation) or data loss.

Web interface software for application site built by using a combination of client side scripting (JavaScript) as an interpretation of real-time programming, server side scripting (PHP) as an API (Application Programming Interface) and Scalable Vector Graphic (SVG). This combination produces programming called PHP-AJAX (Asynchronous JavaScript and XML) to SVG format chart plotter. AJAX is a technique to control the use of JavaScript in communication with the server and then refresh (update) the existing data in a web page without undergoes a refresh an entire Web page as in the usual method.

III. RESULTS AND DISCUSSION

It is very important to make sure that the proposed instrumentation system work well by carry out some tests. Purposes of test is to knows functionality and performance of the sensors, DAQ, and overall of the instrumentation system.

To determine performance of the seismic sensors, it was tested by comparing the sensor output with a commercial geophone. In this case, a comercial geophone placed nearby the developed seismic sensors, then performed vibration test and record both output signal of the sensors. Figure 10 shows signal of the sensor and the geophone. In this figure, upper side signal i.e a small amplitude signal is output of the developed sensor, whereas lower side signal i.e. a large amplitude signal is output of the geophone. Result of experiment show that both signals have same trend. The difference amplitude is caused by deferent amplification in signal conditioning circuit, and it can be handed by chosing suitable amplification. So, from these experimental results can be concluded that the developed sensor has been functioning properly.

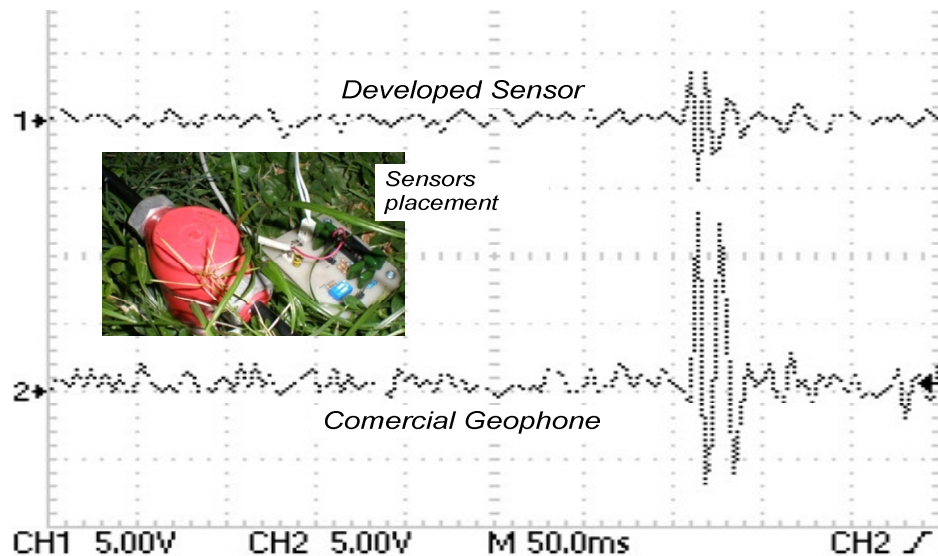


Figure 10. Comparison between developed sensor and geophone

Validation test of the temperature and humidity sensors is carried out by using standard laboratory equipment for calibration i.e. HHF11. Results of calibration are given in Figure 11 for temperature sensor and Figure 12 for humidity sensor. In Figure 11, it appears both sensors have same temperature value for several conditions by $R^2 = 0.99$. In figure 12, there are differences in measurement results between the two sensors especially in high humidity values. This is due to the SHT11 humidity was measured by calculating the compensation of room temperature. However, these differences need to be further analyzed, even necessary recommendations for using other types of humidity sensors.

Validation test of the wind speed sensor is carried out also by using HHF11 calibrator. Calibration is performed by comparing the rotation speed of anemometer cup (developed device) with speed of the calibrator. Graph of calibration is given in Figure 13. The resulting equation is $y = 2.302x$, with $R^2 = 0.99$, will be used to convert the rotational speed of the anemometer cup to wind speed value. For several experimental data, the error of the resulting wind speed sensor is $\pm 5\%$ of HHF11. While the measurement of wind direction by using developed sensors, the results of measurement are given in Table 1.

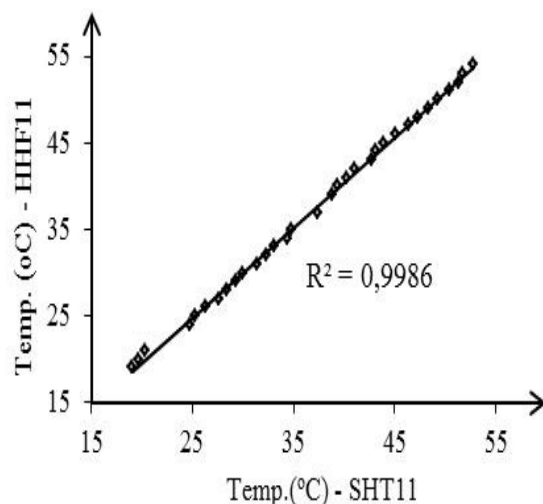


Figure 11. Temperature sensor – calibration

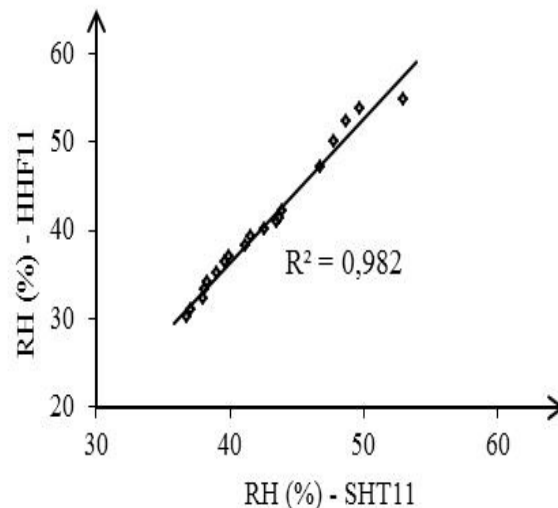


Figure 12. Humidity sensor - calibration

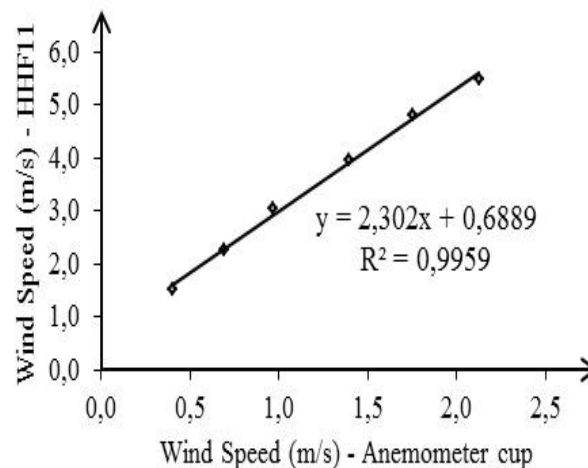


Figure 13. Wind speed sensor – calibration

Table 1. Determination of wind direction

Input (Digital Code)	Direction (°) from North	Input (Digital Code)	Direction (°) from North
0000	0 - 22,5	1000	180 - 202,5
0001	22,5 - 45	1001	202,5 - 225
0010	45 - 67,5	1010	225 - 247,5
0011	67,5 - 90	1011	247,5 - 270
0100	90 - 112,5	1100	270 - 292,5
0101	112,5 - 135	1101	292,5 - 315
0110	135 - 157,5	1110	315 - 337,5
0111	157,5 - 180	1111	337,5 - 360

Furthermore, in case of RTU-MTU communication procedure, it is necessary to do test performance of the telemetry system. For that, seismic signal from sensor in the RTU recorded by using digital oscilloscope, and for the same time it's transmit wirelessly to the MTU by YS-320H RF-transceiver. Figure 14 shows result of the test. Image (a) is the transmitted signal (RTU), and image (b) is the received signal (MTU). It can be seen that both signal have same shape, it's mean that the telemetry system function properly.

Performance test of web based system purposes to find out whether the signal pattern of the graph in MTU's PC (PC's server) corresponds to the pattern shown in the browser client computer. Result of the test shown in Figure 15. It can be seen that is a comparative look between MTU's PC and web browser. However, the signal on the browser client has small delay compare to MTU's PC. This occurs due to distance and wide network connections owned by the PC's client to the server (MTU). In addition, real-time system is defined as a system that not only oriented towards results (outputs) issued, but also required for the system can work well within a certain time as needed.

Finally, Figure 16 shows block diagram of the overall instrumentation system for real time volcano monitoring. The system was designed low-cost in budget and as simple as possible, with the purpose of facilitating the modification if desired, and can be integrated with other systems to improve the quality and quantity of monitoring. According to the some tests that have been done on the overall system, the system is functioning properly. However, field validation still needed to determine performance of the instrumentation system in the real application.

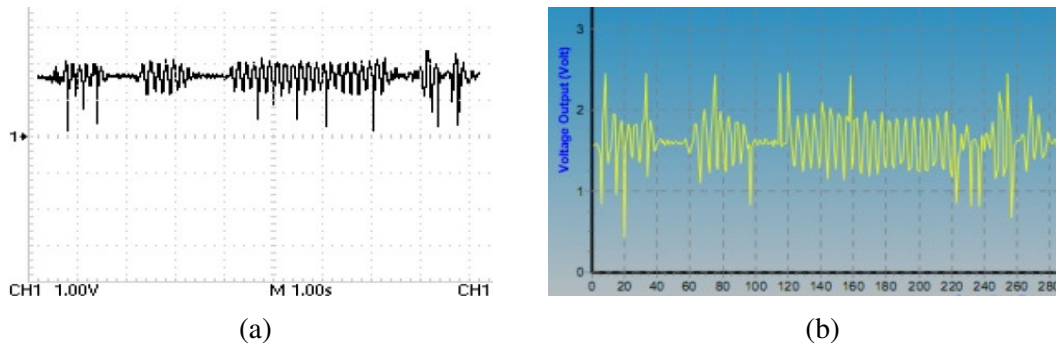


Figure 14. Performance test of the telemetry system (a) transmitting signal, (b) receiving signal

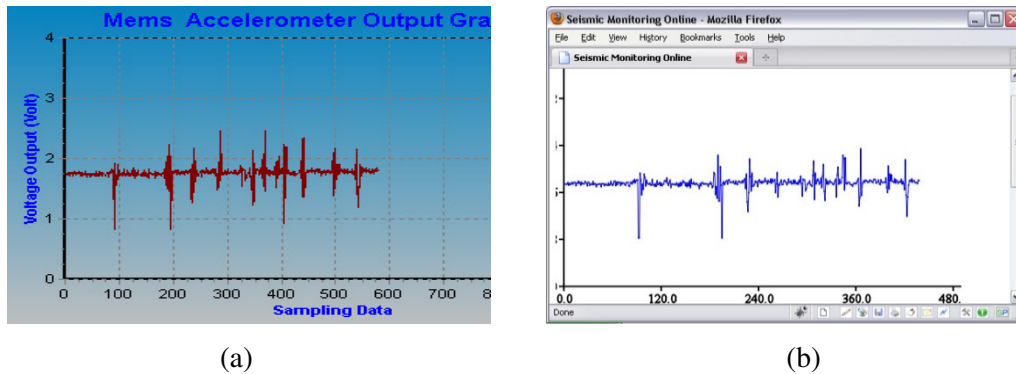


Figure 15. Performance test of the web based application (a) MTU's PC (b) browser client's PC

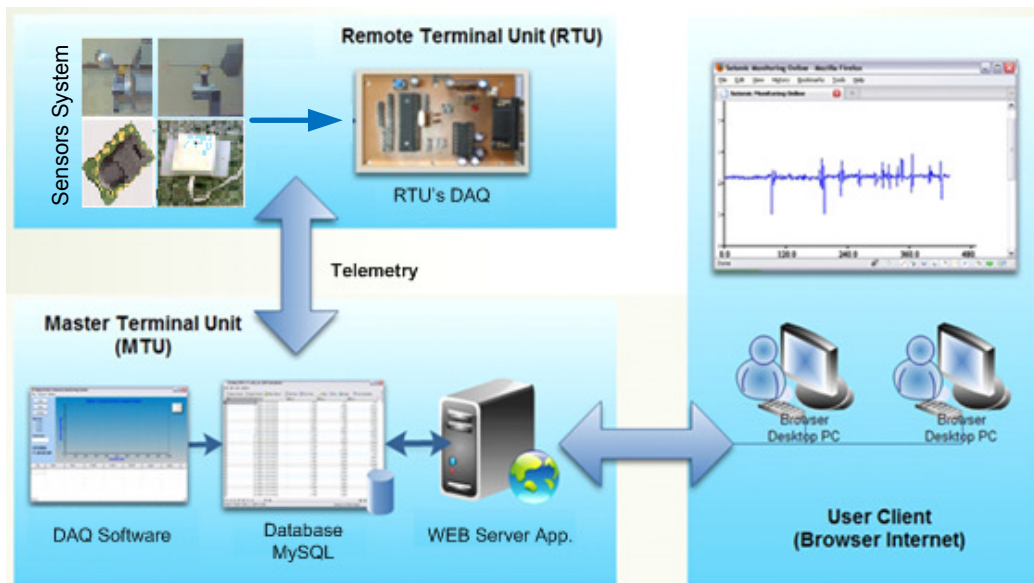


Figure 16. Block diagram of the overall instrumentation system

IV. CONCLUSIONS

In this research, a simple and inexpensive instrumentation system for online volcano monitoring has been developed. Some projects have been successfully tested: (1) seismic sensors and weather sensors, (2) data acquisition and telemetry systems, and (3) logger data and software for web-based monitoring systems. The prototype of seismic sensors offers good results; although still need to do field validation. Weather sensors i.e. air temperature sensor, air humidity sensor, wind speed sensor and wind direction sensor also offer good results. Telemetry system for RTU to MTU wireless communications has been functioning properly. For internet based monitoring system, has been

successfully designed Web-based software applications using MySQL. The monitoring data of seismic activity and the surrounding weather can be recorded and analysed for scientific purposes or mitigation of natural disasters.

ACKNOWLEDGMENTS

The research is financed by The Directorate General of Higher Education, Ministry of National Education, Republic of Indonesia, on the scheme of “*Hibah Penelitian Strategis Nasional*” through DIPA UB No. 0174.0/023-04.2/XV/2009, and “*Hibah Kompetensi*” No. 342.4/UN10.21/PG/2011.

REFERENCES

- [1]. McNutt S. (1996). *Seismic monitoring and eruption forecasting of volcanoes: A review of the state of the art and case histories*. In Scarpa and Tilling, *Monitoring and Mitigation of Volcano Hazards*, Springer-Verlag: Berlin, Heidelberg, pp. 99-146.
- [2]. http://vulcan.wr.usgs.gov/Volcanoes/Indonesia/Maps/map_indonesia_volcanoes.html
- [3]. McGuire B; Kilburn C.R.J.; Murray J. (1995). *Monitoring Active Volcanoes*, UCL Press Limited, London, pp.421.
- [4]. Panagiotopoulos D; Dimitriadis I.; Vamvakaris D. *Seismic Monitoring*. Institute for the Study and Monitoring of the Santorini Volcano. Available on <http://ismosav.santorini.net>
- [5]. Albarbar A.; Mekid S.; Starr A.; Pietruszkiewicz R. (2008). Suitability of MEMS Accelerometers for Condition Monitoring: An experimental study. *Sensors*, 8, 784-799.
- [6]. Rahim I.A.; Miskam, M.A.; Sidek, O.; Zaharudin, S.A.; Zainol, M.Z.; Mohd, S.K.K. (2009). Development of a Vibration Measuring Unit Using a Micro-electromechanical System Accelerometer for Machine Condition Monitoring. *European Journal of Scientific Research*, 35, 1, 150-158.
- [7]. Aizawa T; Kimura T; Matsuoka T; Takeda T; Asano Y. (2008). Application of MEMS accelerometer to geophysics, *International Journal of the JCRM*, 4, 2, pp.1-4.
- [8]. Pascale A. (2009). *Using Micro-ElectroMechanical Systems (MEMS) accelerometers for earthquake monitoring*. Environmental Systems & Services Pty Ltd, Hawthorn, Australia. Available on <http://www.esands.com>
- [9]. Gurrieri S; Madonia P; Giudice G; Inguaggiato S. (2007). Contemporary total dissolved gas pressure and soil temperature anomalies recorded at Stromboli volcano (Italy). *Geophysical Research Letters*, 34 pp 5.
- [10]. Vougioukalakis GE; Fytikas M; Kolios N. *Thermal Monitoring*. Institute for the Study and Monitoring of the Santorini Volcano. Available on <http://ismosav.santorini.net>
- [11]. Wright R; Flynn LP; Garbeil H; Andrew JL; Harris; Eric Pilger. (2004). *MODVOLC: near-real-time thermal monitoring of global volcanism, Hawaii*. Institute of Geophysics and Planetology, University of Hawaii, Honolulu, HI 96822, USA
- [12]. Al-Dhaher. (2001). Integrating hardware and software for the development of microcontroller-based systems. *Microprocessors and Microsystems*, 25, 317-328.
- [13]. Pal, S; Rakshit A. (2004). Development of network capable smart transducer interfaces for traditional sensors and actuators. *Sensor and Actuator A*, 112, 381-387.
- [14]. Flammini A; Ferrari P; Sisinni E; Marioli D; Taroni A. (2003). Sensor integration in industrial environment: from fieldbus to web sensors. *Computer Standards & Interfaces* 25, 183-194.
- [15]. Bentley, J.P. (1995). *Principles of Measurement System*, 3th.ed. Prentice Hall.
- [16]. MMA7260Q, Three-Axis Low g Acceleration Sensor. Available on <http://www.freescale.com>.
- [17]. SHT1x/SHT7x Humidity & Temperature Sensor. Available on <http://www.sensirion.com>.
- [18]. PIC16F876 datasheet. Available on <http://www.microchips.com>.
- [19]. YS-320H, 5W RF-transceiver, wireless data modem. Available on <http://szyishi.bossgoo.com>

Authors

Didik R. Santoso, received B.Sc. degree from Brawijaya University (UB), Malang-Indonesia (1992); M.Sc. degree from Gadjah-Mada University, Yogyakarta-Indonesia (1997), both in Instrumentation Physics; Dr.Eng. degree in Structural Safety System from Hiroshima University, Hiroshima-Japan (2005). At present, he is a senior lecturer and researcher in the Department of Physics, UB. His research field is the smart sensors and instrumentation systems.



Sukir Maryanto, received B.Sc. degree in Physics from Brawijaya University, Malang-Indonesia (1995); M.Sc. degree in Geophysics from Gadjah-Mada University, Yogyakarta-Indonesia (2000); Ph.D degree in Volcano Seismology from Kyoto University, Kyoto-Japan



(2007). At present, he is a senior lecturer and researcher in the Department of Physics, UB. His research field is the Volcano Seismology.

AY Ponco Wardoyo, received B.Sc. degree in Physics from Gadjah-Mada University, Yogyakarta-Indonesia (1988); M.Sc. degree in Astro Physics from University of Tasmania, Hobart-Australia (1995); Ph.D degree in Environmental Physics from Queensland Univ. of Tech., Brisbane-Australia (2007). At present, he is a senior lecturer and researcher in the Department of Physics, UB. His research field is the Environmental Monitoring.



INTENSIFIED ELGAMAL CRYPTOSYSTEM (IEC)

Prashant Sharma¹, Amit Kumar Gupta², Sonal Sharma³

^{1&3}M Tech Student, Sri Balaji College of Engineering & Technology, Jaipur, Rajasthan

²Assistant Professor, Sri Balaji College of Engineering & Technology, Jaipur, Rajasthan

ABSTRACT

In 1985 a powerful and public-key scheme was produced by ElGamal. El-Kassar et al. and El-Kassar and Haraty modified the ElGamal public-key encryption scheme from the domain of natural integers, \mathbb{Z} , to two principal ideal domains, namely the domain of Gaussian integers, $\mathbb{Z}[i]$, and domain of the rings of polynomials over finite fields $F[x]$, by extending arithmetic needed for the modifications to these domains. In this paper we modified the classical ElGamal encryption/decryption algorithm which is based on the difficulty of discrete logarithm problem where it is straight forward to raise numbers to large powers but it is much harder to do the inverse computation of the discrete logarithm. Now, there are so many algorithms available for solving the discrete logarithm problem of small size numbers in a reasonable time. Two of the most popular algorithm is the baby-step/ giant-step algorithm and the pollard's rho algorithm. So to improve security, we proposed a Intensified ElGamal Cryptosystem (IEC) to enhance the security for encrypting long messages and also secure against mathematical and brute-force attack as well as Low-Modulus and Known-Plaintext attack on ElGamal cryptosystem.. The security of this algorithm is based on the difficulty of solving the discrete logarithm problem and Integer factorization problem. This paper also presents comparison between IEC and ElGamal cryptosystem in respect of security and complexity.

KEYWORDS: Encryption, Decryption, Public key, Private Key, Security, ElGamal parameters, Rabin, IFP, DLP.

I. INTRODUCTION

The importance of cryptography system and techniques is becoming a fundamental issue for large sector of society. The use of encryption is so important for both storing and transmitting the data. In order to secure this digital data, strong cryptography techniques are required.

The importance of encryption is increasing rapidly due to the growing traffic on the public network like Internet. Private persons, companies, government agencies and other organizations use encryption techniques in order to safely communicate with partners and customers [1].

Cryptography is the study of mathematical techniques related to aspects of information security which aims to provide the following services [2]:-

- Confidentiality: preventing the unauthorized reception of the message.
- Authentication: verifying the message's origin.
- Data Integrity: establishing that a received message has not been altered.
- Non-Repudiation: This is a service which prevents an *entity* from denying previous commitments or actions.

The encryption are mainly divided into two types; 1) block cipher which split the plaintext into fixed length blocks and then operate on each block separately to produce a corresponding ciphertext block. 2) Stream cipher which operates on the plaintext bit by bit (or character by character) rather than on plaintext block.

The cryptosystem can be one of two types; the symmetric cryptosystem and asymmetric cryptosystem. Symmetric key cryptosystem uses the same key for both encrypting and decrypting the data, this kind is also known as secret key encryption. On the other hand, asymmetric cryptosystem uses two different keys: the first one is the public key used to encrypt the data and the other one is the private key used to decrypt the data, asymmetric cryptosystem is also known as public key encryption [3].

The conceptual difference between symmetric and asymmetric cryptography/encryption are based on how these systems keep a secret. In symmetric key cryptography, the secret must be shared between two persons. In asymmetric cryptography, the secret is personal (unshared); each person creates and keeps his or her own secret [4].

1.1 The Integer Factorization Problem (IFP)

The integer factorization problem is the following: given a composite number n that is the product of two large prime numbers p and q , find p and q .

While finding large prime numbers is a relatively easy task, the problem of factoring the product of two such numbers is considered computationally intractable if the primes are carefully selected. Based on the difficulty of this problem, Rabin and Williams developed Rabin cryptosystem [5].

1.2 The discrete logarithm problem (DLP)

The discrete logarithm problem is the following: given a prime p , a generator α of \mathbb{Z}_p , and a non-zero element $\beta \in \mathbb{Z}_p$, find the unique integer r , $0 \leq r \leq p-2$, such that $\beta \equiv \alpha^r \pmod{p}$. The integer r is called the discrete logarithm of β to the base α . ElGamal scheme is based on discrete logarithm problem [5].

1.3 Rabin Cryptosystem

The Rabin cryptosystem is example of an asymmetric cryptosystem. It was created by Michael O. Rabin in 1979. Rabin's work has theoretical importance because it provided the first provable security for public-key cryptosystems. It can be proven that recovering the entire plaintext from the ciphertext has same complexity as factoring large numbers [2].

Rabin's algorithm is stated as follows:

- Each user chooses two primes p and q each equal to 3 modulo 4, and forms the product $n = p \times q$.
- Public key: the number n .
- Private key: the numbers p and q .
- Encryption: to encrypt a message m , form the ciphertext $c = m^2 \pmod{n}$.
- Decryption: given ciphertext c , use the formulas below to calculate the four square roots modulo n of c : m_1, m_2, m_3 , and m_4 . One of them is the original message m .

In this paper, we compare and evaluate the ElGamal cryptosystem and Intensified ElGamal cryptosystem by implementing and running them on a computer. We investigate the issues of complexity, efficiency and reliability by running the algorithm with different sets of values. Moreover, comparisons will be done between these two algorithms given the same data as input.

The rest of paper is organized as follows- section 2, describes the ElGamal cryptosystem which depends on the discrete logarithm problem and security of ElGamal cryptosystem. In section 3, we present our modified algorithm. In section 4, we compare both the algorithm- ElGamal and Intensified ElGamal cryptosystem. In section 5, describe the security analysis of our modified algorithm. A conclusion is shown in section 6.

II. ELGAMAL CRYPTOSYSTEM

The classical Elgamal encryption-decryption scheme is one of the most popular and widely used public-key cryptosystem. It is based on the difficulty of the discrete logarithm problem for finite fields. The ElGamal Cryptosystem consists of three steps [2]:

Step 1) Generating the key

1. Choose a large prime number p , such that $(p - 1) / 2$ is prime, too. N is the number of bits of p .
2. Choose the base $\alpha < p$.
3. Choose the private key $a < p$.
4. Compute $\beta = \alpha^a \pmod{p}$

5. Publish p , alpha and beta as public key.

Step 2) Encryption of a message

1. Split the plaintext into blocks of $N - 1$ bits (fill them up if necessary).
2. Choose a secret random number k with $\gcd(k, p - 1) = 1$
3. Compute for every block x the ciphertext
 $e(x, k) = (y_1, y_2)$, where $y_1 = \alpha^k \pmod{p}$ and $y_2 = \beta^k x \pmod{p}$. y_1 and y_2 are blocks of the length N of ciphertext.

Step 3) Decryption of a message

1. Split the ciphertext in blocks of N bits
2. For two successive blocks y_1 and y_2 solve
 $y_1^a x = y_2 \pmod{p}$ for x . Thus
 $d(y_1, y_2) = x = y_2 (y_1^a)^{-1} \pmod{p}$
 is the wanted block of plaintext.

2.1 Security of ElGamal Cryptosystem

The security of ElGamal cryptosystem is also broken by two attacks based on low modulus and known-plain text attacks [4].

A) *Low-Modulus Attack*: If the value of p is not large enough, then broker can use efficient algorithms to solve the discrete logarithm problem to find alpha or private key “a”. It means that the security of ElGamal cryptosystem depends on the infeasibility of solving a discrete logarithm problem with a very large module [4].

B) *Known-Plain text Attack*: If sender use same random exponent k , to encrypt two plaintexts P and P' , broker discover P' if he knows P . So we have to use fresh value of k for different plaintext [4].

III. OUR SCHEME (IEC)

In this section, we propose an efficient algorithm for enciphering a long plaintext. Our proposed algorithm (Intensified Elgamal Cryptosystem (IEC)) is based on both ElGamal cryptosystem [2] and Rabin cryptosystem [2].

Our cryptosystem algorithm consists of three steps:

Step 1) Generating the key

1. Choose a large prime number p , such that $(p - 1) / 2$ is prime, too. N is the number of bits of p .
2. Choose the base alpha $< p$.
3. Choose the private key $a < p$.
4. Compute $\beta = \alpha^a \pmod{p}$
5. Generate two large random (and distinct) primes r and s , each roughly the same size and both primes when divided by 4 give remainder 3
6. Compute $n = r * s$.
7. Use the extended Euclidean algorithm to find integers A and B satisfying $A * r + B * s = 1$. Note that A and B can be computed once and for all during the key generation stage.
8. Publish p , n , alpha and beta as public key and kept secret A , B , a , r , s as private key.

Step 2) Encryption of a message

1. Split the plaintext into blocks of $N - 1$ bits (fill them up if necessary).
2. Choose a secret random number k with $\gcd(k, p - 1) = 1$. (There should be different k for different plain text message)
3. Compute for every block x the ciphertext
 $e(x, k) = (y_1, y_2)$, where $y_1 = \alpha^k \pmod{p}$ and $y_2 = \beta^k x \pmod{p}$. y_1 and y_2 are blocks of the length N of ciphertext.
4. Append ten consecutive five (“5”) digits at the end of y_1 and at the end of y_2 then add y_1 with y_2 as $m = y_1 + (10 \text{ consecutive “5” digits}) + y_2 + (10 \text{ consecutive “5” digits})$.
5. Compute $c = m^2 \pmod{n}$.
6. c is the cipher text message.

Step 3) Decryption of a message

To recover cipher text m from c , one should do the following:

1. Compute $t_1 = c^{(r+1)/4} \bmod r$.
2. Compute $t_2 = c^{(s+1)/4} \bmod s$.
3. Compute $x = (A*r*t_2 + B*s*t_1) \bmod n$.
4. Compute $y = (A*r*t_2 - B*s*t_1) \bmod n$.
5. Using x and y , we get four messages m_1, m_2, m_3 and m_4 as follows:
6. $m_1 = x, m_2 = -x, m_3 = y$, and $m_4 = -y$.
7. One of the four is the original message m and we can identify the original message by using the redundant consecutive bits which is appended at the last of original message m .
8. Now by using the message m , we find the plain text as follows
9. We divide the message m into two sub blocks y_1 and y_2 by using the redundant consecutive bits added at the end of y_1 .
10. For two successive blocks y_1 and y_2 solve
Compute $x = y_2 (y_1^{-1}) \bmod p$.
11. x is the wanted block of plaintext.

IV. COMPARISON BETWEEN IEC AND ELGAMAL CRYPTOSYSTEM

4.1 Simulation Results

For the simulation purpose, IEC cryptosystem is implemented as a user-friendly GUI. This GUI application is implemented using Java BigInteger library functions [6]. In this application, one can either enter the prime numbers or can specify the bit length of the prime numbers to generate automatically. BigInteger library provides operations for modular arithmetic, GCD calculation, primarily testing, prime generation, bit manipulation, and a few other miscellaneous operations. The simulation of the algorithm, implemented in JAVA [6], running on a 2.8 GHz P-IV Processor and 512 MB RAM, using a 1000 characters long message for encryption/decryption. In this section, we investigate the issues of complexity, efficiency and reliability by running the program with different sets of data.

The comparison between both algorithms, are analyzed by giving the same data as input and changing only one parameter at a time while keeping the other parameter unchanged. The PKC algorithms (ElGamal & IEC) have some important parameters affecting its level of security and speed [1]. The complexity of decomposing the modules into its factors is increased by increasing the module length. This also increases the length of private key and hence difficulty to detect the key. The algorithm depends mainly on the *prime number* whose value depends on the number of bits used to generate this prime. It also depends on the *private key*, the *input message* and the chunk size (*cut length*) of the block message [1]. The key generation, encryption, decryption time is depends on the speed of the processor and the RAM. The results are shown through the graph on the basis of the readings in the Table 1 for the Module Size 2048 bit. Actually the module size is varying from 256 bit to 2048 bit, but we are showing results only for Module Size 2048 bit.

4.1.1 Changing the length of the prime number (p) length/module size (n)

The value of the prime number (p), plays an important role in forming the values of the public key (p , α , β) and hence the encrypted message pair (y_1, y_2). This prime number is a function of the number of bits used to generate it. Changing this bit number is directly reflected on the length of the other parameter as shown in Table 1.

The length of the module size (n) is directly proportional to its generating number of bits since the maximum value of (p). The length of encrypted message (y_1, y_2) also increase as a function of the prime number (p). As a result, increasing the n -bit length provides more security. On other hand by increasing the n -bit length increases the values of key generation time, encryption time and decryption time as shown in Figure 1 and 2.

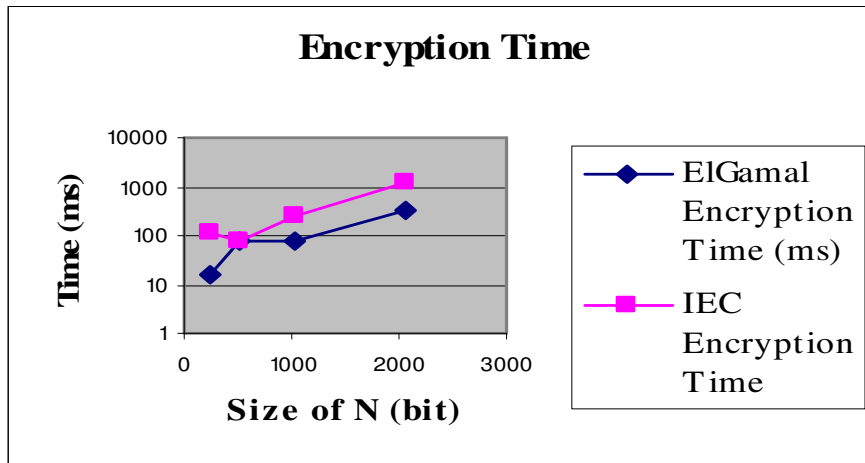


Figure1. Modulus (N) Size (bit) v/s Encryption time (ms), taking size of Private Key (a) 32 bit and Chunk Size 64 bit.

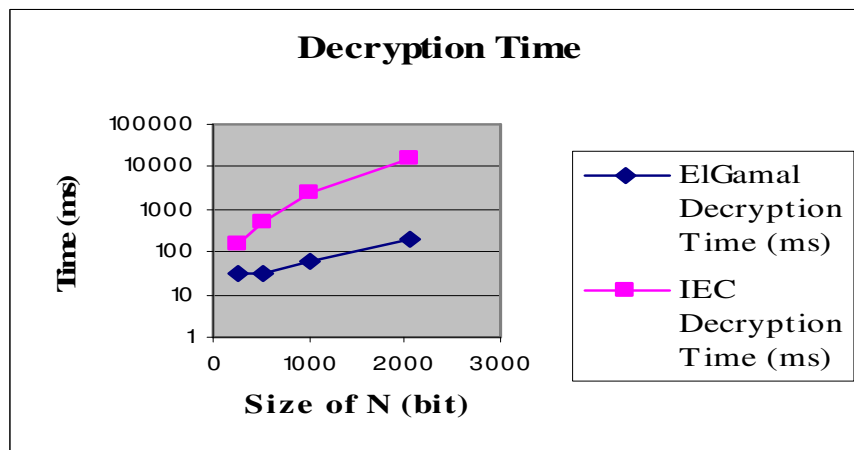


Figure2. Modulus (N) Size (bit) v/s Decryption time (ms), taking size of Private Key (a) 32 bit and Chunk Size 64 bit

4.1.2 Changing the private key length

The private key plays an important role in generating the public key and ciphertext values. To study the performance of private key, the value of module size (n) and chunk size is fixed and the value of private key is changed. The Figure 3 and 4 shows the speed of encryption and decryption with variable private key.

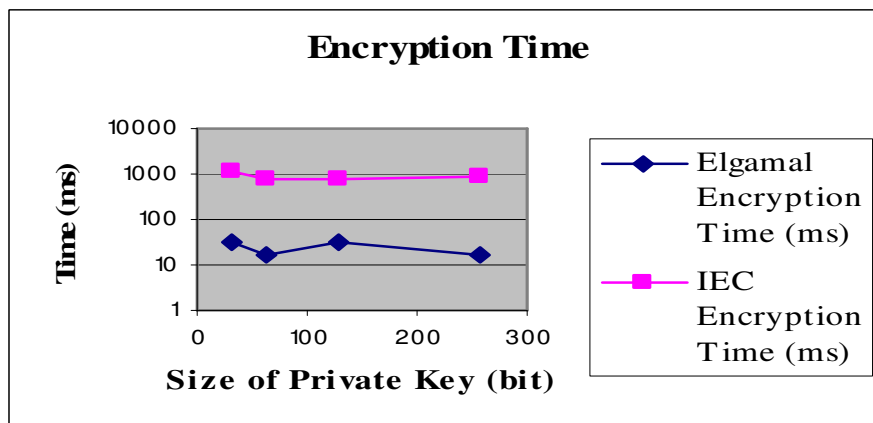


Figure3. Private Key (a) size v/s Encryption time, taking Modulus (N) size 2048 bit and chunk size 64 bit.

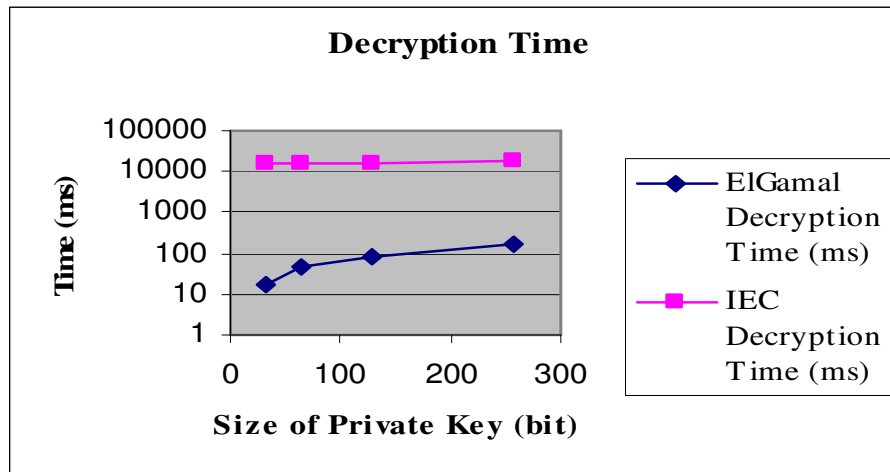


Figure4: Private Key (a) size v/s Decryption time, taking Modulus (N) size 2048 bit and chunk size 64 bit.

4.1.3 Changing the cut length (or chunk size)

To study the performance of chunk size the value of module size (n) and private key kept unchanged and change the value of chunk size. The Figure 5 and 6 shows the speed of encryption and decryption with variable chunk size is continuously decrease.

Table 1: Effect of changing the Chunk Size and Private Key with the constant module size (N) 2048 bit

Module (N) size (bit)	p size (bit)	Private Key "a" size (bit)	Chunk Size (bit)	ELGAMAL CRYPTOSYSTEM				IEC			
				key generation time (ms)	Encryption time (ms)	Decryption time (ms)	Total Execution time(ms)	Key Generation time (ms)	Encryption time (ms)	Decryption time (ms)	Total Execution time(ms)
2048	1024	32	64	15	313	187	500	17531	1141	15609	16750
2048	1024	32	128	15	63	94	157	17531	141	6031	6172
2048	1024	32	256	15	46	47	93	17531	47	2766	2813
2048	1024	32	512	15	31	16	47	17531	16	1344	1360
2048	1024	64	64	16	250	359	609	10078	797	16062	16859
2048	1024	64	128	16	63	172	235	10078	94	6140	6234
2048	1024	64	256	16	32	93	125	10078	94	2813	2907
2048	1024	64	512	16	16	47	63	10078	15	1407	1422
2048	1024	128	64	16	281	687	968	15625	797	16219	17016
2048	1024	128	128	16	63	375	438	15625	141	6297	6438
2048	1024	128	256	16	31	157	188	15625	47	2860	2907
2048	1024	128	512	16	31	78	109	15625	15	1391	1406
2048	1024	256	64	16	250	1265	1515	108859	875	17406	18281
2048	1024	256	128	16	63	672	735	108859	125	6735	6860
2048	1024	256	256	16	31	328	359	108859	47	3031	3078
2048	1024	256	512	16	16	156	172	108859	16	1531	1547

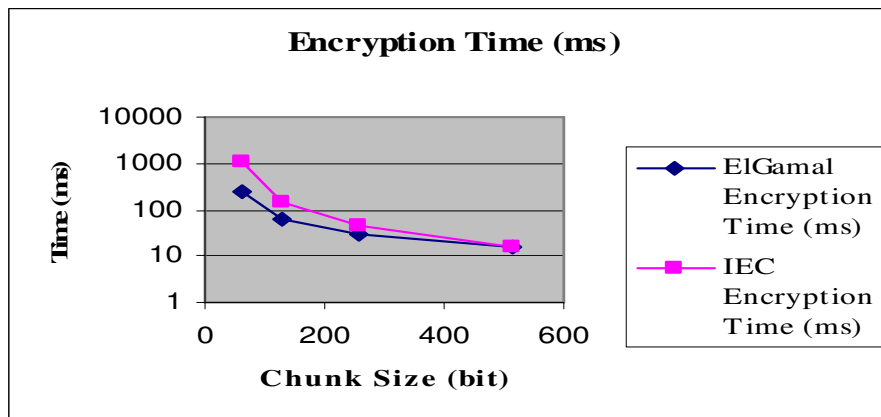


Figure5. Chunk size v/s Encryption time, taking size of Modulus (N) 2048 bit and size of Private Key 32 bit.

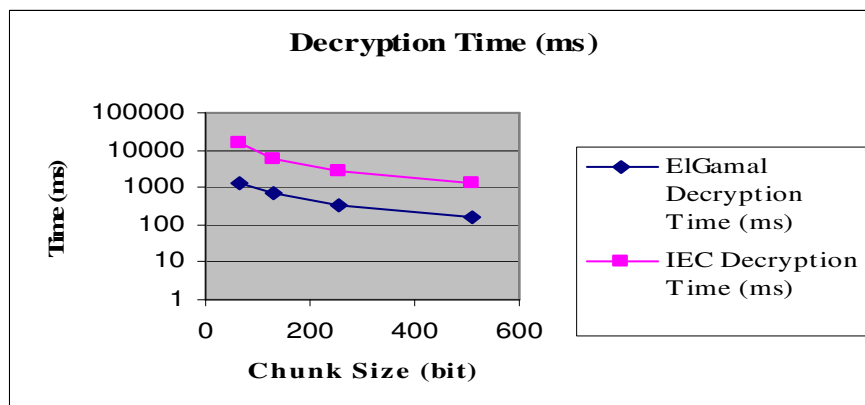


Figure6. Chunk size v/s Decryption time, taking size of Modulus (N) 2048 bit and size of Private Key 32 bit.

4.2 Complexity analysis of IEC cryptosystem

Let n be the module used in the ElGamal cryptosystem. ElGamal encryption, decryption, each take $2n$ time scalar point multiplication and $2n$ times multiplication operation. Rabin encryption takes one modular multiplication and one modular squaring. Rabin decryption is slower than encryption; it means the complexity of Rabin system is as factoring a large number n into its two prime factors p and q . So in this way, computational complexity of IEC cryptosystem depends on $2n$ times scalar point multiplication and factoring of two large prime numbers p and q . The simulation results of both the algorithms shows that the execution time of IEC is 2 times more than ElGamal.

V. SECURITY ANALYSIS

Two possible approaches to attack the IEC / ElGamal algorithms are as follows [6, 7]:

- *Brute force*: This involves trying all possible private keys.
- *Mathematical attacks*: These are based on factoring the product of large primes such as factor n into its prime factors p and q respectively.

Estimated resources needed to factor a number within one year are as follows [8].

Table 2: Recourses needed to factor a number within one year

Size of number (Bits)	PCs	Memory
430	1	128 MB
760	215,000	4 GB
1020	342×10^6	170 GB
1620	1.6×10^{15}	120 TB

The security of our scheme is based on both Integer Factorization Problem (IFP) [5] and Discrete Logarithm Problem (DLP) [5], it is very difficult for an illegal user to compute the secret key 'a' from the equation $\beta = \alpha^a \pmod{p}$. It is also difficult for an intruder to obtain the system-generated random numbers r and s directly from equation $A*r + B*s = 1$ of IEC Algorithm. The difficulty relies on the complexity of computing discrete logarithm over finite fields. In *integer factorization problem (IFP)* finding large prime numbers is a relatively easy task, the problem of factoring the product of two such numbers is considered computationally intractable if the primes are carefully selected. For long-term security, 1024-bit or larger modulus p should be used in *discrete logarithm problem (DLP)* and *integer factorization problem (IFP)*.

The discrete logarithm problem can be solved by *naïve* [2], *baby-step/giant-step* [2] and *Pollard's rho* algorithm [2]. Because of these algorithms, the security of ElGamal cryptosystem is broken by the *Low-Modulus* [4] and *Known-Plaintext attack* [4]. As the IEC cryptosystem uses dual modulus, so for breaking this one have to factor both the modulus. That's why our cryptosystem provides the far better security against the *naïve*, *baby-step/giant-step* and *Pollard's rho* algorithm and attacks on ElGamal cryptosystem.

VI. CONCLUSION

The ElGamal cryptosystem is secured against passive attack but not against adaptive attack. Efficient algorithms for the discrete logarithm problem would render the ElGamal Cryptosystem insecure. So to improve the security, this paper presents an Intensified ElGamal Cryptosystem (IEC). IEC algorithm is based on the combination of factorization of large number and the discrete logarithm problem. So if one want to break the IEC, one have to factor large numbers into its prime factors and also have to solve discrete logarithms problems. This paper also shows comparisons among IEC and ElGamal cryptosystems in terms of security & performance. The disadvantage of new cryptosystem is that, unlike ElGamal it can not be used for authentication as it is based on the one way function. Another disadvantage is the slow down of execution process as compare to ElGamal. But it is clear from the simulation results that it is more secure than the ElGamal algorithm and our cryptosystem provides the far better security against the *naïve*, *baby-step/giant-step* and *Pollard's rho* algorithm and attacks on ElGamal cryptosystem.

REFERENCES

- [1] Allam Mousa, "Security and Performance of ElGamal Encryption Parameter," Journal of Applied Sciences 5, Asian Network for Scientific Information, 883-886, 2005.
- [2] Alfred J Menezes, Paul C van Oorschot, Scot A Vanstone, Handbook of Applied Cryptography, CRC Press, 1997.
- [3] Wenbo Mano, "Modern Cryptography Theory and Practice," Prentice Hall.
- [4] Behrouz A. Forouzan, Debdeep Mukhopadhyay, "Cryptography and Network Security" 2nd Edition, TMH.
- [5] <http://www.certicom.com>, "The Elliptic Curve Cryptosystem," September 1997, (dated 02-04-2010)
- [6] Neal R. Wagner, "The Laws of Cryptography with Java Code", Technical Report, 2003, pages 78-112
- [7] William Stallings, "Cryptography and Network Security", ISBN 81-7758-011-6, Pearson Education, Third Edition, pages 42-62, 121-144, 253-297.
- [8] CHUK, Chinese university (2009), "RSA Algorithm security and Complexity", Retrieved from [http://www.cse.cuhk.edu.hk/~phwl/mt/public/archives/old/ceg50 10/rsa.pdf](http://www.cse.cuhk.edu.hk/~phwl/mt/public/archives/old/ceg50%20rsa.pdf) (dated 04-04-2011)
- [9] J H Moore, Protocol failures in Cryptosystems, Contemporary Cryptology, The science of Information Integrity, Ed. G J Simmons, 541-558, IEEE Press 1992.
- [10] H E Rose, "A Course in Number Theory, Second Addition", Oxford Science Publications, 1994.
- [11] El-Kassar, A.N. and Haratly Ramzi, ElGamal Public Key Cryptosystem Using Reducible Polynomials over a Finite Field, Proceedings of the 13th International Conference on Intelligent & Adaptive Systems and Software Engineering (IASSE-2004). Nice, France. July 2004.
- [12] El-Kassar, A. N., Chihadi H., and Zentout D. Quotient rings of polynomials over finite fields with cyclic group of units, Proceedings of the International Conference on Research Trends in Science and Technology, pp. 257-266, 2002.
- [13] Menezes, A. J., Van Oorschot, and Vanstone, P.C. S.A. Handbook of Applied Cryptography, CRC press, 1997.
- [14] B. Schneizer, Applied Cryptography, New York: John Wiley, 1994.
- [15] <http://www.certicom.com>, "The Elliptic Curve Cryptosystem," September 1997, (dated 02-04-2010)

[16] J H Moore, Protocol failures in Cryptosystems, Contemporary Cryptology, The science of Information Integrity, Ed. G J Simmons, 541-558, IEEE Press 1992.

[17] Diffie, M Hellman, "New Directions in Cryptography", IEEE Transactions on Information Theory, Vol 22, 1976.

Author's Biography

Prashant Sharma was born on 04 June 1985. He is the M.Tech. Student in SBCET, Jaipur (Rajasthan). He has completed B.E.(I.T.) in 2007 from University of Rajasthan, Jaipur.



Amit Kumar Gupta was born on 11 Nov 1981. He is working as an Assistant Professor in SBCET, Jaipur. He has 4.5 year teaching experience. He has completed B.E. (I.T.) and MTech. (CSE).



Sonal Sharma was born on 23 May 1984. She is M.Tech. Student in SBCET, Jaipur (Rajasthan). She has completed B.E. (I.T.) in 2006 from University of Rajasthan, Jaipur.



A ZIGZAG-DELTA PHASE-SHIFTING TRANSFORMER AND THREE-LEG VSC BASED DSTATCOM FOR POWER QUALITY IMPROVEMENT

R.Revathi¹ and J.Ramprabu²

¹PG Student, Department of Electrical and Electronics Engineering, Kumaraguru College of Technology, Coimbatore-641 049, India

²Assistant Professor, Department of Electrical and Electronics Engineering, Kumaraguru College of Technology, Coimbatore-641 049, India

ABSTRACT

In this paper, a new three-phase four-wire distribution static compensator (DSTATCOM) based on a Phase-Shifting transformer (PST) and a three-leg voltage source converter (VSC) is proposed for power quality improvement. The Zigzag Phase-Shifting transformer mitigates the circulating power flows in interconnected utilities and the three-leg VSC compensates harmonic current, reactive power and balances the load. The principle of PST is to take harmonics generated from separate sources, shift one source of harmonics 180° with respect to the other and then combine them together; this will result in cancellation. This type of transformer has patented built-in electromagnetic technology designed to remove high neutral current and the most harmful harmonics from the 3rd through 21st. The interesting aspect of such phase shifters is that despite their low MVA capacity, by controlling the phase shift, they exercise a significant real-power control. The insulated gate bipolar transistor (IGBT) based VSC is supported by a capacitor and is controlled for the required compensation of the load current. The dc bus voltage of the VSC is regulated during varying load conditions. The performance of the three-phase four-wire DSTATCOM is validated using MATLAB software with its Simulink and power system blockset toolboxes.

KEYWORDS: Distribution Static Compensator (DSTATCOM), Neutral Current Compensation, Power Quality Improvement, Phase-Shifting Transformer (PST), Voltage Source Converter (VSC)

I. INTRODUCTION

Three-phase four-wire distribution systems are facing severe power quality problems such as poor voltage regulation, high reactive power and harmonics current burden, load unbalancing, excessive neutral current, etc. Three-phase four-wire distribution systems are used in commercial buildings, office buildings, hospitals, etc. Most of the loads in these locations are nonlinear loads and are mostly unbalanced loads in the distribution system. This creates excessive neutral current both of fundamental and harmonic frequency, and the neutral conductor gets overloaded. The voltage regulation is also poor in the distribution system due to the unplanned expansion and the installation of different types of loads in the existing distribution system. In order to control the power quality problems, many standards are proposed, such as the IEEE-519 standard [6].

There are mitigation techniques for power quality problems in the distribution system and the group of devices is known by the generic name of custom power devices (CPDs) [1]. The distribution static compensator (DSTATCOM) is a shunt-connected CPD capable of compensating power quality problems in the load current. Some of the topologies of DSTATCOM for three-phase four-wire

system for the mitigation of neutral current along with power quality compensation in the source current are four-leg voltage source converter (VSC), three single-phase VSCs, three-leg VSC with split capacitors [3], three-leg VSC with zigzag transformer [7]-[9] and three-leg VSC with neutral terminal at the positive or negative of dc bus [10]. The voltage regulation in the distribution feeder is improved by installing a shunt compensator [11]. There are many control schemes reported in the literature for control of shunt active compensators such as instantaneous reactive power theory, power balance theory, synchronous reference frame theory, symmetrical components based, etc. [12], [13]. The synchronous reference frame theory [12] is used for the control of the proposed DSTATCOM. In this paper, a new topology of DSTATCOM is proposed for a three-phase four-wire distribution system, which is based on three-leg VSC and a Zigzag Phase-Shifting transformer (PST). Harmonic cancellation is performed with harmonic cancelling transformers also known as phase-shifting transformers. A harmonic cancelling transformer is a relatively new power quality product for mitigating harmonic problems in electrical distribution systems. This type of transformer has patented built-in electromagnetic technology designed to remove high neutral current and the most harmful harmonics from the 3rd through 21st. The technique used in these transformers is call "low zero phase sequencing and phase shifting". These transformers can be used to treat existing harmonics in buildings or facilities. This same application can be designed into new construction to prevent future harmonics problems. The insulated gate bipolar transistor (IGBT) based VSC is self-supported with a dc bus capacitor and is controlled for the required compensation of the load current.

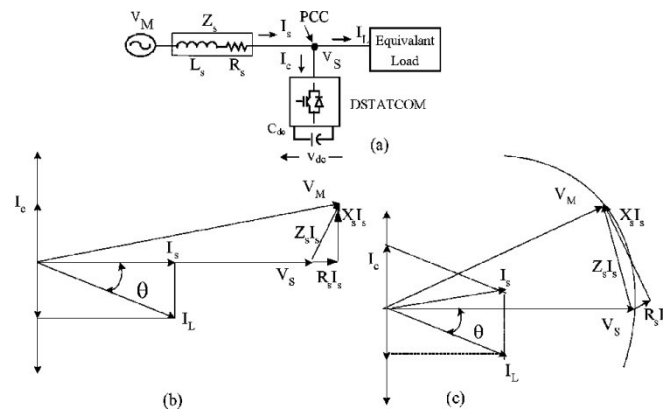


Fig. 1. (a) Single-line diagram of DSTATCOM system. (b) Phasor diagram for UPF operation. (c) ZVR operation.

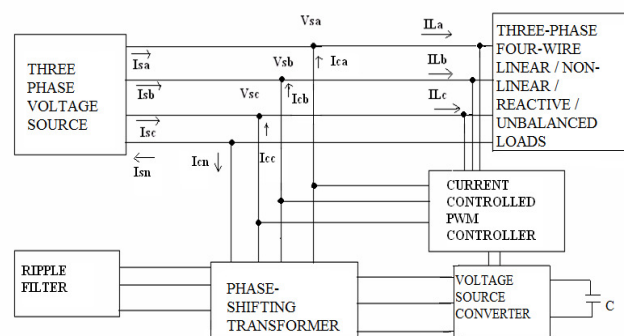


Fig. 2. Block diagram of the proposed three-leg VSC with Phase-Shifting transformer based DSTATCOM connected in distribution system.

The DSTATCOM is designed and simulated using MATLAB software with its Simulink and power system blockset (PSB) toolboxes for power factor correction and voltage regulation along with neutral current compensation, harmonic elimination, and load balancing with linear loads as well as nonlinear loads.

II. SYSTEM CONFIGURATION AND DESIGN

Fig. 1(a) shows the single-line diagram of the shunt-connected DSTATCOM based distribution system. The dc capacitor connected at the dc bus of the converter acts as an energy buffer and establishes a dc voltage for the normal operation of the DSTATCOM system. The DSTATCOM can be operated for reactive power compensation for power factor correction or voltage regulation. Fig. 1(b) shows the phasor diagram for the unity power factor operation. The DSTATCOM injects a current I_C such that the source current is only I_S , and this is in-phase with voltage. The voltage regulation operation of DSTATCOM is depicted in the phasor diagram of Fig. 1(b). The DSTATCOM injects a current I_C such that the voltage at the load (V_S) is equal to the source voltage (V_M).

The proposed DSTATCOM consists of a three-leg VSC and a Phase-Shifting transformer is shown in Fig. 2, where the Phase-Shifting transformer is connected in Zigzag – delta configuration. The Zigzag Phase-Shifting Transformer block implements a three-phase transformer with a primary winding connected in a zigzag configuration and a configurable secondary winding. The model uses three single-phase, three-winding transformers. The primary winding connects the windings 1 and 2 of the single-phase transformers in a zigzag configuration. The secondary winding connects the windings 3 of the single-phase transformers in delta configuration.

A three-leg VSC is used as an active shunt compensator along with a Phase-shifting transformer, as shown in Fig.2, and this topology has six IGBTs, three ac inductors, and one dc capacitor. The required compensation to be provided by the DSTATCOM decides the rating of the VSC components. The VSC is designed for compensating a reactive power of 12 kvar, as decided from the load details. The selection of interfacing inductor, dc capacitor, and the ripple filter are given in the following sections.

2.1. DC Capacitor Voltage

The minimum dc bus voltage of VSC of DSTATCOM should be greater than twice the peak of the phase voltage of the system [17]. The dc bus voltage is calculated as

$$V_{dc} = \frac{2\sqrt{2}V_{LL}}{\sqrt{3}m} \quad (1)$$

where m is the modulation index and is considered as 1, and V_{LL} is the ac line output voltage of DSTATCOM. Thus, V_{dc} is obtained as 677.69 V for V_{LL} of 415 V and is selected as 700 V.

2.2. DC Bus Capacitor

The value of dc capacitor (C_{dc}) of VSC depends on the instantaneous energy available to the DSTATCOM during transients [17]. The principle of energy conservation is applied as

$$\frac{1}{2}C_{dc}[(V_{dc}^2) - (V_{dc1}^2)] = 3V(\alpha I)t \quad (2)$$

where V_{dc} is the reference dc voltage and V_{dc1} is the minimum voltage level of dc bus, α is the overloading factor, V is the phase voltage, I is the phase current, and t is the time by which the dc bus voltage is to be recovered.

Considering the minimum voltage level of the dc bus, $V_{dc1} = 690$ V, $V_{dc} = 700$ V, $V = 239.60$ V, $I = 27.82$ A, $t = 350$ μ s, $\alpha = 1.2$, the calculated value of C_{dc} is 2600 μ F and is selected as 3000 μ F.

2.3. AC Inductor

The selection of the ac inductance (L_f) of VSC depends on the current ripple $i_{cr,p-p}$, switching frequency f_s , dc bus voltage (V_{dc}), and L_f is given as [17]

$$L_f = \frac{\sqrt{3}m V_{dc}}{12\alpha f_s i_{cr(p-p)}} \quad (3)$$

where m is the modulation index and α is the overload factor. Considering, $i_{cr,p-p} = 5\%$, $f_s = 10$ kHz, $m = 1$, $V_{dc} = 700$ V, $\alpha = 1.2$, the L_f value is calculated to be 2.44 mH.

2.4. Ripple Filter

A low-pass first-order filter tuned at half the switching frequency is used to filter the high-frequency noise from the voltage at the PCC. Considering a low impedance of 8.1Ω for the harmonic voltage at a frequency of 5 kHz, the ripple filter capacitor is designed as $C_f = 5 \mu\text{F}$. A series resistance (R_f) of 5Ω is included in series with the capacitor (C_f). The impedance is found to be 637Ω at fundamental frequency, which is sufficiently large, and hence, the ripple filter draws negligible fundamental current.

2.5. Design of the Phase-Shifting Transformer

A basic three-phase Zigzag-delta transformer consists of a Zigzag-connected primary and a delta-connected secondary. The zigzag part is accomplished by winding half of the primary turns of one phase of the transformer on one leg of the three-phase transformer, with the other half of the primary turns on an adjacent phase. The schematic diagram of the basic Zigzag-delta transformer is shown in fig.3 (a) below:

From the figure 3(a) we can see that A and A' are wound on the same leg similarly B and B' are so wound as well as C and C'. In operation, the harmonics current flow from the harmonics generating loads into the transformer primary windings. With all triplen-harmonics current are in phase with each other, by vector analysis, the positive and negative flux interaction in the zigzag is "cancelling" these triplen-harmonics. Hence there is reduced harmonics current flowing back into the primary and system.

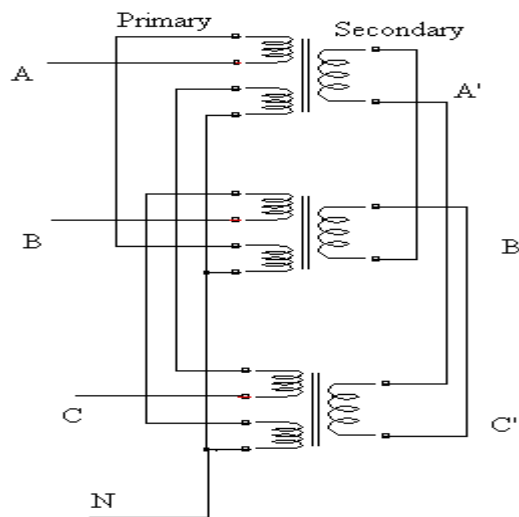


Fig. 3. (a) Zigzag-Delta Transformer Connection.

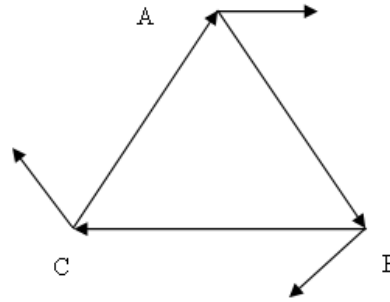


Fig. 3. (b) Phasor diagram.

Delta connected secondary winding transformer will force the triplen harmonics to circulate at its secondary winding and eliminated balanced triplen harmonics. From the phasor diagram Fig. 1(b), it can be found that the voltage across the transformer's winding is $\frac{1}{\sqrt{3}}$ of the phase voltage of the three-phase four-wire distribution power system.

III. CONTROL OF DSTATCOM

The control approaches available for the generation of reference source currents for the control of VSC of DSTATCOM for three-phase four-wire system are instantaneous reactive power theory (IRPT), synchronous reference frame theory (SRFT), unity power factor (UPF) based, instantaneous symmetrical components based, etc. [12], [13]. The SRFT is used in this investigation for the control of the DSTATCOM. A block diagram of the control scheme is shown in Fig. 4. The load currents (i_{La} , i_{Lb} , i_{Lc}), the PCC voltages (V_{Sa} , V_{Sb} , V_{Sc}), and dc bus voltage (V_{dc}) of DSTATCOM are sensed as feedback signals. The load currents from the a-b-c frame are first converted to the α - β -o frame and then to the d-q-o frame using

$$\begin{bmatrix} \dot{i}_{Lq} \\ \dot{i}_{Ld} \\ \dot{i}_{Lo} \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \cos \theta & \cos(\theta - \frac{2\pi}{3}) & \cos(\theta + \frac{2\pi}{3}) \\ \sin \theta & \sin(\theta - \frac{2\pi}{3}) & \sin(\theta + \frac{2\pi}{3}) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \dot{i}_{La} \\ \dot{i}_{Lb} \\ \dot{i}_{Lc} \end{bmatrix} \quad (4)$$

where $\cos \theta$ and $\sin \theta$ are obtained using a three-phase phase-locked loop (PLL). A PLL signal is obtained from terminal voltages for generation of fundamental unit vectors [18] for conversion of sensed currents to the d-q-o reference frame. The SRF controller extracts dc quantities by a low-pass filter, and hence, the non-dc quantities (harmonics) are separated from the reference signal. The d-axis and q-axis currents consist of fundamental and harmonic components as

$$\dot{i}_{Ld} = \dot{i}_{d\ dc} + \dot{i}_{d\ ac} \quad (5)$$

$$\dot{i}_{Lq} = \dot{i}_{q\ dc} + \dot{i}_{q\ ac} \quad (6)$$

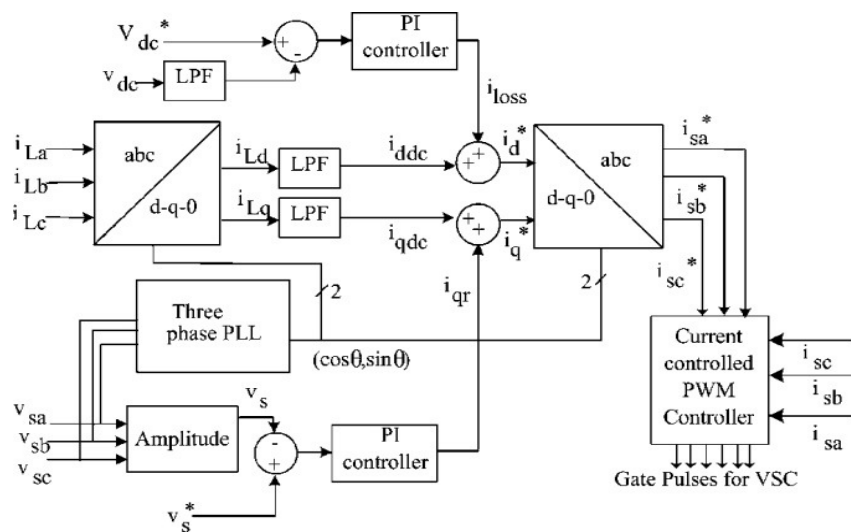


Fig. 4. Control algorithm for the three-leg-VSC-based DSTATCOM in a three-phase four-wire system.

3.1. UPF Operation of DSTATCOM

The control strategy for reactive power compensation for UPF operation considers that the source must deliver the mean value of the direct-axis component of the load current along with the active power component current for maintaining the dc bus and meeting the losses (i_{loss}) in DSTATCOM. The output of the proportional-integral (PI) controller at the dc bus voltage of DSTATCOM is considered as the current (i_{loss}) for meeting its losses

$$\dot{\mathbf{i}}_{loss(n)} = \dot{\mathbf{i}}_{loss(n-1)} + \mathbf{K}_{pd} (V_{dc(n)} - V_{dc(n-1)}) + \mathbf{K}_{id} V_{dc(n)} \quad (7)$$

where $V_{de(n)} = V_{dc}^* - V_{dc(n)}$ is the error between the reference (V_{dc}^*) sensed (V_{dc}) dc voltages at the nth sampling instant. K_{pd} and K_{id} are the proportional and integral gains of the dc bus voltage PI controller. The reference source current is therefore

$$\dot{\mathbf{i}}_d^* = \dot{\mathbf{i}}_{d\ dc} + \dot{\mathbf{i}}_{loss} \quad (8)$$

The reference source current must be in phase with the voltage at the PCC but with no zero-sequence component. It is therefore obtained by the following reverse Park's transformation with $\dot{\mathbf{i}}_d^*$ as in (12) and $\dot{\mathbf{i}}_q^*$ and $\dot{\mathbf{i}}_0^*$ as zero

$$\begin{bmatrix} \dot{\mathbf{i}}_a^* \\ \dot{\mathbf{i}}_b^* \\ \dot{\mathbf{i}}_c^* \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 1 \\ \cos(\theta - \frac{2\pi}{3}) & \sin(\theta - \frac{2\pi}{3}) & 1 \\ \cos(\theta + \frac{2\pi}{3}) & \sin(\theta + \frac{2\pi}{3}) & 1 \end{bmatrix} \begin{bmatrix} \dot{\mathbf{i}}_d^* \\ \dot{\mathbf{i}}_q^* \\ \dot{\mathbf{i}}_0^* \end{bmatrix} \quad (9)$$

3.2. Zero-Voltage Regulation (ZVR) Operation of DSTATCOM

The compensating strategy for ZVR operation considers that the source must deliver the same direct-axis component $\dot{\mathbf{i}}_d^*$, as mentioned in (12) along with the sum of quadrature-axis current ($\dot{\mathbf{i}}_{q\ dc}$) and the component obtained from the PI controller ($\dot{\mathbf{i}}_{qr}$) used for regulating the voltage at PCC. The amplitude of ac terminal voltage (V_s) at the PCC is controlled to its reference voltage (V_s^*) using the PI

controller. The output of PI controller is considered as the reactive component of current (i_{qr}) for zero-voltage regulation of ac voltage at PCC. The amplitude of ac voltage (V_s) at PCC is calculated from the ac voltages (V_{sa} , V_{sb} , V_{sc}) as

$$V_s = (2/3)^{1/2} (V_{sa}^2 + V_{sb}^2 + V_{sc}^2)^{1/2} \quad (10)$$

Then, a PI controller is used to regulate this voltage to a reference value as

$$i_{qr(n)} = i_{qr(n-1)} + K_{pq} (V_{te(n)} - V_{te(n-1)}) + K_{iq} V_{te(n)} \quad (11)$$

where $V_{te(n)} = V_s^* - V_{s(n)}$ denotes the error between reference (V_s^*) and actual ($V_{s(n)}$) terminal voltage amplitudes at the n^{th} sampling instant. K_{pq} and K_{iq} are the proportional and integral gains of the dc bus voltage PI controller. The reference source quadrature-axis current is

$$i_q^* = i_{qdc} + i_{qr} \quad (12)$$

The reference source current is obtained by reverse Park's transformation using (13) with i_d^* as in (12) and i_q^* as in (16) and i_0^* as zero.

3.3. Computation of Controller Gains

The gains of the controllers are obtained using the Ziegler-Nichols step response technique [19]. A step input of amplitude (U) is applied and the output response of the dc bus voltage is obtained for the open-loop system. The maximum gradient (G) and the point at which the line of maximum gradient crosses the time axis (T) are computed. The gains of the controller are computed using the following equations:

$$K_p = \left| \frac{1.2U}{GT} \right| \quad (13)$$

$$K_i = \left| \frac{0.6U}{GT^2} \right| \quad (14)$$

The gain values for both the PI controllers are computed.

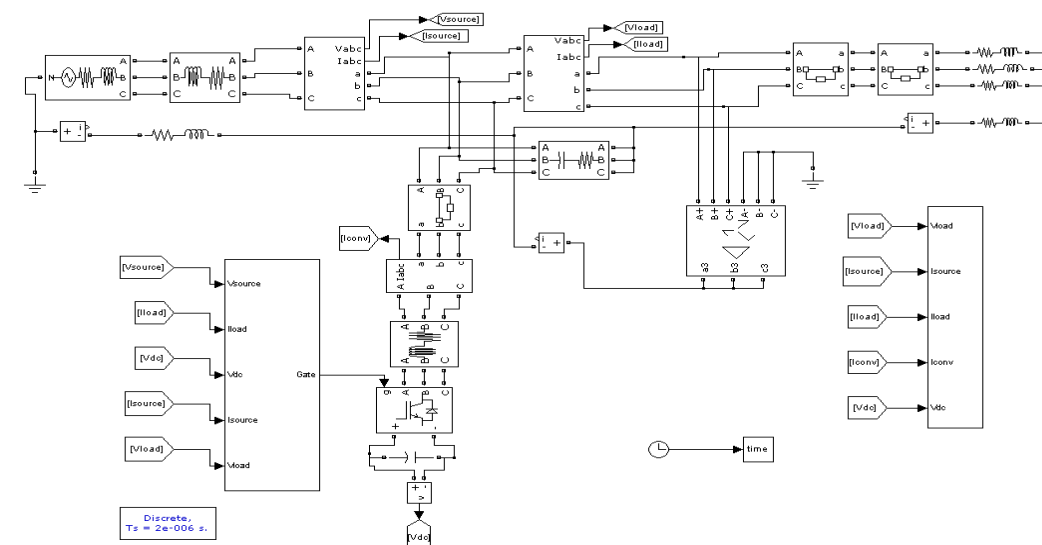


Fig. 5. MATLAB model of the Zigzag-delta configuration of the Phase-Shifting transformer and the three-eg

VSC based DSTATCOM connected system.

3.4. Current-Controlled Pulse Width Modulation (PWM) Generator

In a current controller, the sensed and reference source currents are compared and a proportional controller is used for amplifying current error in each phase before comparing with a triangular carrier signal to generate the gating signals for six IGBT switches of VSC of DSTATCOM.

IV. MODELING AND SIMULATION

The three-leg VSC and the Phase-Shifting Transformer based DSTATCOM connected to a three-phase four-wire system is modeled and simulated using the MATLAB with its Simulink and PSBs. The ripple filter is connected to the DSTATCOM for filtering the ripple in the PCC voltage. The MATLAB-based model of the three-phase four-wire DSTATCOM is shown in the Fig.5. The Phase-Shifting transformer in parallel to the load, the three-phase source, and the shunt-connected three-leg Voltage Source Converter are connected as shown in the Fig.5. The available model of Zigzag Phase-Shifting transformer is used for modeling the Zigzag-delta configuration of the Phase-Shifting transformer.

The control algorithm for the DSTATCOM is also modeled in MATLAB. The reference source currents are derived from the sensed PCC voltages (V_{sa} , V_{sb} , V_{sc}), load currents (i_{La} , i_{Lb} , i_{Lc}) and the dc bus voltage of DSTATCOM (V_{dc}). A PWM current controller is used over the reference and sensed source currents to generate the gating signals for the IGBTs of the VSC of the DSTATCOM.

V. RESULTS AND DISCUSSION

The performance of the Phase-Shifting Transformer and three-leg-VSC based three-phase four-wire DSTATCOM is demonstrated for power flow control, power factor correction and voltage regulation along with harmonic reduction, load balancing, and neutral current compensation. The developed model is analyzed under linear load the results are discussed shortly.

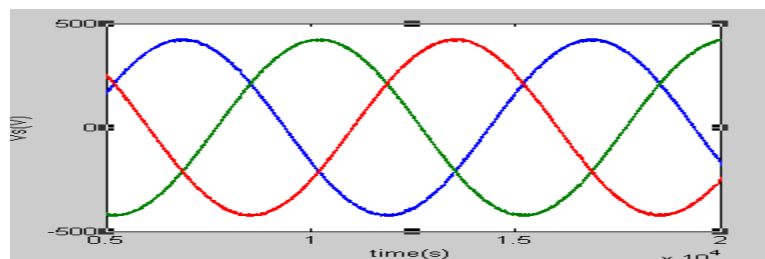


Fig. 6.(a) PCC Voltage (V)

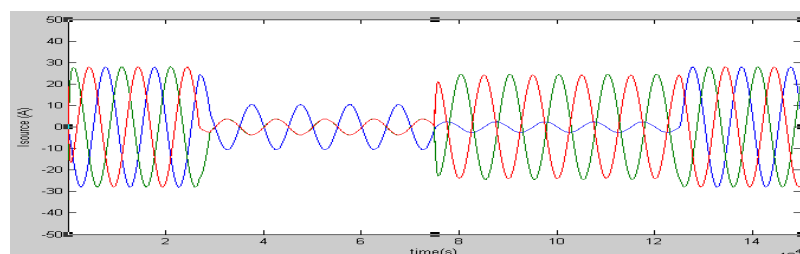


Fig. 6.(b) Source current (A)

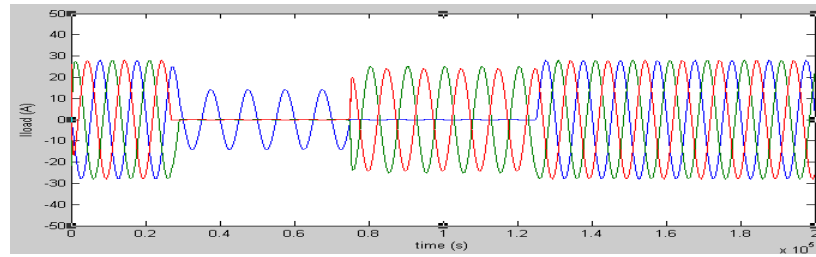


Fig. 6.(c) Load current (A)

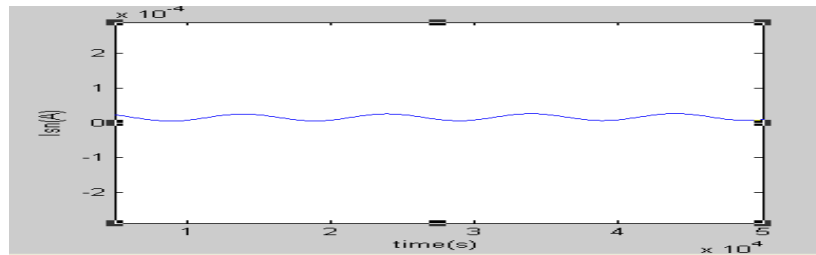


Fig. 6.(d) Source-neutral current (A)

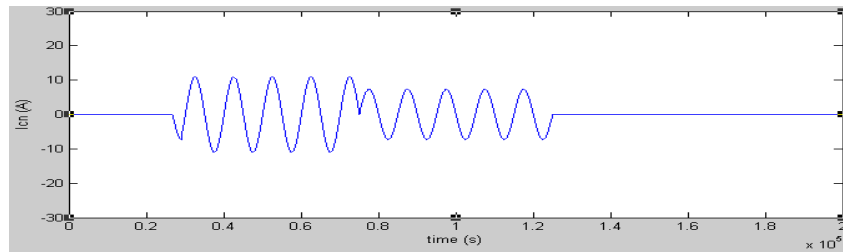


Fig. 6.(e) Compensator-neutral current (A)

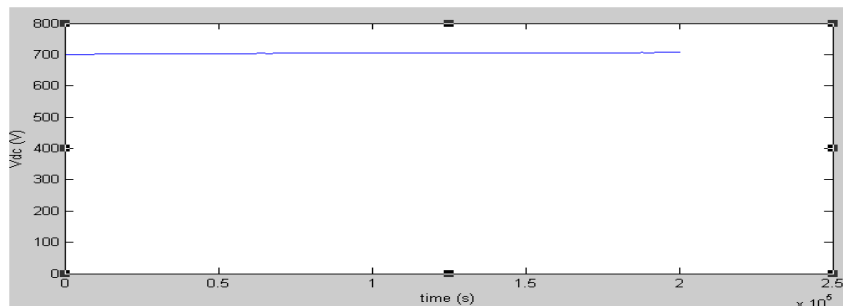


Fig. 6.(f) DC Bus Voltage (V)

5.1. Performance of DSTATCOM with linear load for Neutral Current Compensation, Load Balancing, UPF Operation and Zero-Voltage Regulation Operation

The dynamic performance of the DSTATCOM under linear lagging power factor unbalanced load condition is shown in Fig. 6. At 0.6 s, the load is changed to two-phase load and to single-phase load at 0.7 s. These loads are applied again at 0.8 and 0.9 s, respectively. The PCC voltages (V_s), source currents (i_s), load currents (i_L), compensator currents (i_c), source-neutral current (i_{sn}), load-neutral current (i_{Ln}), compensator-neutral current (i_{cn}), dc bus voltage (V_{dc}), and amplitude of voltage (V_s) at PCC are also depicted in Fig. 6. The source-neutral current is observed as nearly zero, and this verifies the proper compensation. The reactive power is compensated for power factor correction, and the source currents are balanced and sinusoidal. It is also observed that the dc bus voltage of DSTATCOM is able to maintain close to the reference value under all disturbances. Its total harmonic

distortion is low. The amplitude of PCC voltage is maintained at the reference value under various load disturbances, which shows the ZVR mode of operation of DSTATCOM.

5.2. Experimental Demonstration of the Performance of Zigzag Phase-Shifting Transformer

The proposed topology of DSTATCOM consists of combined operation of three-leg VSC and a Zigzag Phase-Shifting transformer. As the performance of a three-leg VSC and a zig-zag transformer is studied in [9], the Zigzag-delta configuration of the Phase-Shifting transformer is analyzed for the compensation of neutral current. A prototype of the Zigzag-delta configuration of the Phase-Shifting transformer is developed in the laboratory and the neutral current compensation technique is tested for linear and nonlinear loads. The Phase-Shifting transformer is tested for neutral current compensation under unbalanced linear loads and balanced/unbalanced nonlinear loads. When the load is nonlinear and balanced, the load-neutral current is observed to be mainly triplen harmonics current, and this is compensated using the Phase-Shifting transformer.

VI. COMPARISON WITH OTHER TECHNIQUES

A three-leg single-phase-VSC-based DSTATCOM [3] requires a total of 12 semiconductor devices, and hence, is not attractive, and the three-leg VSC with split capacitors [3] has the disadvantage of difficulty in maintaining equal dc voltage at two series-connected capacitors. The four-leg-VSC-based DSTATCOM [3] is considered as superior considering the number of switches, complexity, cost, etc. The proposed three-phase-four-wire DSTATCOM is based on a three-leg VSC and a Phase-Shifting transformer. A three-leg VSC with zig-zag and delta configuration of Phase-Shifting transformer has the advantage of using a passive device for neutral current compensation, reduced number of switches, use of readily available three-leg VSC, etc. It is observed that the KVA rating of the PST is less compared to the T-connected transformer. Similarly, comparison with the four-leg converter shows that the number of switches are reduced in the proposed configuration, thus reducing the complexity and cost of the system. The total harmonic distortion of the Zigzag-delta configuration of the Phase-Shifting transformer is less compared to any other techniques. The advantage of the PST is that their low MVA capacity and it improves the stability of the power system. The T-connected transformer requires two single-phase transformers, whereas the zig-zag transformer reported in [9] has three single-phase transformers with a turn's ratio of 1:1. A star/delta transformer is also reported [20] for neutral current compensation and the kVA rating required is higher compared to Phase-Shifting transformer. Table I shows the rating comparison of the three transformer techniques for a given neutral current of 10 A. But the disadvantage of the proposed topology is that it requires the protection of PST.

Table 1. Comparison of Rating of Transformer Connections.

Transformer	Winding Voltage (V)	Winding Current (A)	KVA	Total KVA
T-connected	240/120/120	10	2.4	4.48
	208/208		2.08	
Star/Delta	240/240	10	2.4	7.2
Zigzag Phase-Shifting	120/120	10	1.4	4.2

VII. CONCLUSION

The performance of a new topology of three-phase four-wire DSTATCOM consisting of three-leg VSC with a PST has been demonstrated for reactive power compensation, harmonic elimination, load balancing and mitigating circulating power flows in interconnected utilities. The voltage regulation and power factor correction modes of operation of the DSTATCOM have been observed and are as

expected. The PST improves the transient stability of a power system. It is also observed that the Phase-Shifting transformer helps in damping low frequency power oscillations. The dc bus voltage of the DSTATCOM has been regulated to the reference dc bus voltage under varying loads. The total kilovolt-amperes rating of the PST is lower than the T-connected transformer for reactive power compensation. The PST has verified that it is effective in compensating the reactive power, zero sequence fundamental and harmonic currents. We found that harmonic mitigation techniques using Zigzag-delta transformers are a viable choice for mitigating rich triplen harmonics in four-wire electrical distribution system. As comparison made in term of costs, Zigzag-delta Phase-Shifting transformers show that low cost as to compare with active filter and other types of harmonic mitigation. The proposed approach is much simpler, cheaper and reliable compared to approaches using active filters. The costs comparison with active filter as shown below:

Active Filter – RM1750/kVA

Transformer – RM 300/kVA

Zigzag-delta Phase-Shifting transformers are also significantly reduced neutral current flowing back to the system almost up to 90%.

ACKNOWLEDGEMENTS

The authors wish to thank the family members who have provided full support.

REFERENCES

- [1] A. Ghosh and G. Ledwich, *Power Quality Enhancement Using Custom Power Devices*. London, U.K.: Kluwer, 2002.
- [2] R. C. Dugan, M. F. McGranaghan, and H. W. Beaty, *Electric Power Systems Quality*, 2nd ed. New York: McGraw-Hill, 2006.
- [3] H. Akagi, E. H. Watanabe, and M. Aredes, *Instantaneous Power Theory and Applications to Power Conditioning* Hoboken, NJ: Wiley, 2007.
- [4] A. Moreno-Munoz, *Power Quality: Mitigation Technologies in a Distributed Environment*. London, U.K.: Springer-Verlag, 2007.
- [5] E. F. Fuchs and M. A. S. Mausoum, *Power Quality in Power Systems and Electrical Machines*. London, U.K.: Elsevier, 2008.
- [6] IEEE Recommended Practices and Requirements for Harmonics Control in Electric Power Systems, IEEE Standard 519, 1992.
- [7] L. H. Beverly, R. D. Hance, A. L. Kristalinski, and A. T. Visser, "Method and apparatus for reducing the harmonic currents in alternating current distribution networks," U.S. Patent 5 576 942, Nov. 19, 1996.
- [8] H.-L. Jou, J.-C. Wu, K.-D. Wu, W.-J. Chiang, and Y.-H. Chen, "Analysis of zig-zag transformer applying in the three-phase four-wire distribution power system," *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 1168–1173, Apr. 2005.
- [9] H.-L. Jou, K.-D. Wu, J.-C. Wu, and W.-J. Chiang, "A three-phase four-wire power filter comprising a three-phase three-wire active filter and a zig-zag transformer," *IEEE Trans. Power Electron.*, vol. 23, no. 1, pp. 252–259, Jan. 2008.
- [10] H. L. Jou, K. D. Wu, J. C. Wu, C. H. Li, and M. S. Huang, "Novel power converter topology for three-phase four-wire hybrid power filter," *IET Power Electron.*, vol. 1, no. 1, pp. 164–173, 2008.
- [11] H. Fugita and H. Akagi, "Voltage-regulation performance of a shunt active filter intended for installation on a power distribution system," *IEEE Trans. Power Electron.*, vol. 22, no. 1, pp. 1046–1053, May 2007.
- [12] M. C. Benhabib and S. Saadate, "New control approach for four-wire active power filter based on the use of synchronous reference frame," *Electr. Power Syst. Res.*, vol. 73, no. 3, pp. 353–362, Mar. 2005.

- [13] M. I. Milanes, E. R. Cadaval, and F. B. Gonzalez, "Comparison of control strategies for shunt active power filters in three-phase four-wire systems," *IEEE Trans. Power Electron.*, vol. 22, no. 1, pp. 229–236, Jan. 2007.
- [14] B. A. Cogbill and J. A. Hetrick, "Analysis of T–T connections of two single phase transformers," *IEEE Trans. Power App. Syst.*, vol. PAS-87, no. 2, pp. 388–394, Feb. 1968.
- [15] IEEE Guide for Applications of Transformer Connections in Three-Phase Distribution Systems, IEEE C57.105-1978 (R2008).
- [16] B. Singh, V. Garg, and G. Bhuvaneswari, "A novel T-connected autotransformer-based 18-pulse AC–DC converter for harmonic mitigation in adjustable-speed induction-motor drives," *IEEE Trans. Ind. Electron.*, vol. 54, no. 5, pp. 2500–2511, Oct. 2007.
- [17] B. N. Singh, P. Rastgoufard, B. Singh, A. Chandra, and K. A. Haddad, "Design, simulation and implementation of three pole/four pole topologies for active filters," in *Inst. Electr. Eng. Proc. Electr. Power Appl.*, Jul. 2004, vol. 151, no. 4, pp. 467–476.
- [18] S. Bhattacharya and D. Diwan, "Synchronous frame based controller implementation for a hybrid series active filter system," in *Proc. IEEE Ind. Appl. Soc. Meeting 1995*, pp. 2531–2540.
- [19] J. R. Hendershot and T. J. E. Miller, "Design of Brushless Permanent Magnet Motors," Oxford, U.K.: Magna Physics, 1994.
- [20] J. Arrillaga and R. M. Duke, "Thyristor-controlled quadrature boosting," *Proc. IEE*, v.126, n.6, June 1979, pp. 493–498.
- [21] C. P. Arnold, R. M. Duke and J. Arrillaga, "Transient stability improvement using thyristor controlled quadrature voltage injection," *IEEE Trans.*, v. PAS – 100, n.3, March 1981, pp. 1382–1388.
- [22] R. M. Mathur and R. S. Basati, "A thyristor controlled static phase-shifter for AC power transmission," *IEEE Trans.*, v. PAS – 100, n. 5, May 1981, pp. 2650–2655.
- [23] H. Stemmler and G. Guth, "A thyristor controlled static phase shifter, a new tool for power flow control in AC transmission systems," *Brown Boveri Review*, v. 69, n. 2, 1982, pp. 73–78.
- [24] R. Baker, G. Guth, W. Egli and P. Eglin, "Control algorithm for a static phase shifting transformer to enhance transient and dynamic stability of large power systems," *IEEE Trans.*, v. PAS-101, n.9, September 1982, pp. 3532–3542.

Authors

R. Revathi was born in Erode, India, in 1986. She received the B.E (Electrical and Electronics) degree in Velalar College of Engineering and Technology, Thindal, Erode, India in 2008. During 2008-2010, she worked as a Lecturer in M. Kumarasamy College of Engineering, Karur, India. Now she is doing M.E (Power Electronics and Drives) in Kumaraguru College of Technology, Coimbatore, India. Her areas of interest are Power Electronics and Drives, Power Quality, and Renewable Energy.



J. Ramprabu was born in Coimbatore, India, in 1984. He received the B.E. (Electrical and Electronics) degree from Sri Ramakrishna Engineering College, Coimbatore, India, in 2005. He received the M.E (Applied Electronics) degree in Kumaraguru College of Technology, Coimbatore, India, in 2008. He received the M.B.A (International Business Management) degree from Anna University of Technology, Coimbatore, India, in 2011. During 2005-2006, he worked as a Lecturer in MPNMJ Engineering College, Chennimalai, Erode, India. In 2008, he joined the Department of Electrical and Electronics Engineering, Kumaraguru College of Technology, Coimbatore, as a Lecturer and became an Assistant Professor in 2011. His current research interests include Applied Electronics, Power Quality, and Smart Grid. Mr. Ramprabu is a Member of the Indian Society for Technical Education.



MODELING AND SIMULATION OF NANOSENSOR ARRAYS FOR AUTOMATED DISEASE DETECTION AND DRUG DELIVERY UNIT

S.M. Ushaa¹, Vivek Eswaran²

¹Principal, Siddhartha Institute of Science and Technology, Puttur, Affiliated to Jawaharlal
Nehru Technological University, Anantapur, Andhra Pradesh, India

²Department of Electronics and Communication Engineering, RMK Engineering College
Gummidipundi, TamilNadu, India

ABSTRACT

In this paper, the mathematical models required to describe the functionality of nanodevices have been reviewed. Based on these mathematical models sensor equivalent circuits have been developed. An experimental setup is developed to analyze the characteristics of IS Field Effect Transistor (ISFET), nanowire and nanosphere devices. The impact of geometrical properties on device performance is estimated based on the experimental setup. Settling time and surface analyte concentration graphs obtained using the experimental setup is used in designing a nanobio sensor for disease detection. Based on the test results, a mathematical model has been developed in Matlab to model nanodevices. Three different iterations of sensor models are carried out based on the results obtained curve fitting techniques are adopted to generalize the developed sensor model using Savitzky-Golay Filter (SG Filter). The sensors modeled can be used for automated drug detection and delivery unit. In this paper it is proposed to develop a methodology for integrating biosensor models from nanohub.org and Matlab. The model for nanowire based sensor may be developed using basic principles and can be characterized using experimental setup. Sensor array model consisting of 64 nanowires is proposed to develop to detect prostate cancer. A control unit that triggers the sensor array may be developed and can be used in measuring the concentration of analyte solution. The location of nanowire sensors on the 8 x 8 matrix can be distributed using Gaussian distribution function. A new sensor array consisting of planar sensor and nanowire sensor may be developed to increase the sensitivity of the system in detecting prostate cancer. Expert system based on feed forward neural network architecture may be designed and modeled for ovarian cancer classification. A two layered network consisting of sigmoid transfer function and purelin function may be designed. Optimum weights for the layers are to be identified using Back propagation Levenberg-Marquardt (LM) algorithm. Linear Discriminant Analysis (LDA) and Principle Component Analysis (PCA) techniques are to be incorporated into the expert system that classifies the cancer cells based on significant features extracted. A Proportional-Integral-Derivative (PID) controller may be modeled to control the diffusion pump and monitoring of drug diffusion. The output of expert system should drive the PID control for drug diffusion. Field Programmable Gate Array (FPGA) implementation of neural network architecture and PID controller may be designed and developed for optimizing area, speed and power performances. The modules can be integrated to form the automated disease detection and drug delivery unit.

The developed sensor model for nanowires match the sensor models available from standard resources with less than 2% deviation. Prostate Specific Antigen (PSA) antibody and Deoxyribonucleic Acid (DNA) as biomarkers for detection of prostate cancer based on sensor array is built. The sensor array model achieves 91% efficiency and is superior compared with existing design. The expert system developed correctly classifies ovarian cancer 98% times, with only 2% error. The decision produced by the expert system drives PID controller to activate the diffusion pump, the PID controller has overshoot error of less than 12% with settling time less than 10ms. FPGA implementation of expert system operates at maximum frequency of 293 MHz occupies less than 148

slices when targeted on Virtex 4 FPGA. For real time disease detection and diagnosis, the developed system can be incorporated into a biochip.

KEY WORDS: Nanobio Sensors; Disease Detection; Sensor Modelling; Nanowire; PSA

I. INTRODUCTION

Exhaustive studies and developments in the field of nanotechnology have been carried out and different nanomaterials have been utilized to detect cancer at early stages [Ludwig, j. *et al* (2005)]. Nanomaterials have unique physical, optical and electrical properties that have proven to be useful in sensing. Quantum dots, gold nanoparticles, magnetic nanoparticles, carbon nanotubes, gold nanowires and many other materials have been developed over the years. Nanotechnology has been developing rapidly during the past few years and with this, properties of nanomaterials are being extensively studied and many attempts are made to fabricate appropriate nanomaterials [Catalona, w. (1996), Ushaa eswaran, *et al* (2006)]. Due to their unique optical, magnetic, mechanical, chemical and physical properties that are not shown at the bulk scale, nanomaterials have been used for more sensitive and precise disease detection. For developing a system to detect disease, software modeling is one of the major requirements. Matlab environment is predominantly used for developing software reference models. Various sensor models (electrical and mechanical) are already inbuilt in Matlab and are readily available for development of automotive and mechanical system [Ushaa eswaran *et al* (2009), (2004), Beckett, m. *et al* (1995)]. There are a large number of nanobio sensors that are being used for medical applications in disease detection. There is a need for a mathematical model of nanobio sensor for developing a software reference model in disease detection using Matlab. Thus in this work, we develop a mathematical model for nanowire, that is used for cancer detection. Section II discusses the geometrical and mathematical models of nanowires. Section III discusses the diffusion capture model that is used for modeling nanosensor, section IV presents the experimental setup for simulation of nanowire sensors and design of biosensors. Section V presents the Matlab models developed based on the simulation results obtained and Savitzky-Golay Filter technique to improve the accuracy of sensor models developed.

II. DNA SENSORS

Human genomes have billions of DNA base spheres to sense the DNA sequence. Arrays of sensors are used for genome sensing. Nanobio sensor consists of X-Y array of elements. These elements further consist of pixels called as electronic components [Ludwig, j.(2005)]. Each component is a sensor that can be a nanowire transistor, carbon nanotube, planar transistor etc. Each element has a unique and known DNA sequence (genome) bound to the sensor. As in figure 1, Q1 is one such sequence consisting of ACGAT [Henderson, c. *et al* (1998), Dandachi, n. *et al* (2002) Molina, r. *et al* (2008)] molecule arranged in an order. Each location in the X-Y array has a known sequence attached to it. Figure 1 shows the array of sensors, and the corresponding DNA sequence attached to the sensor.

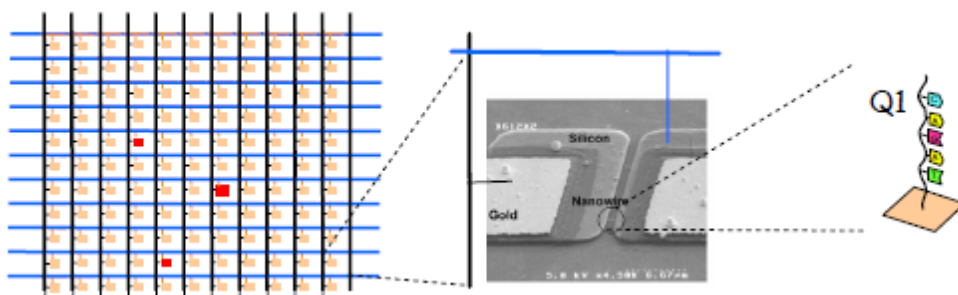


Figure 1 Nanosensor array and DNA sensor [Landman, j. *et al* (1998)]

When an unknown DNA sequence is introduced into the XY array, the unknown sequence finds its conjugate in the XY array and binds with the DNA sequence present on the array as shown in Figure 2. Since the DNA sequence at every location along the XY array is known, the binding of unknown

sequence with known DNA sequence modulates the current in the corresponding element in the XY array.

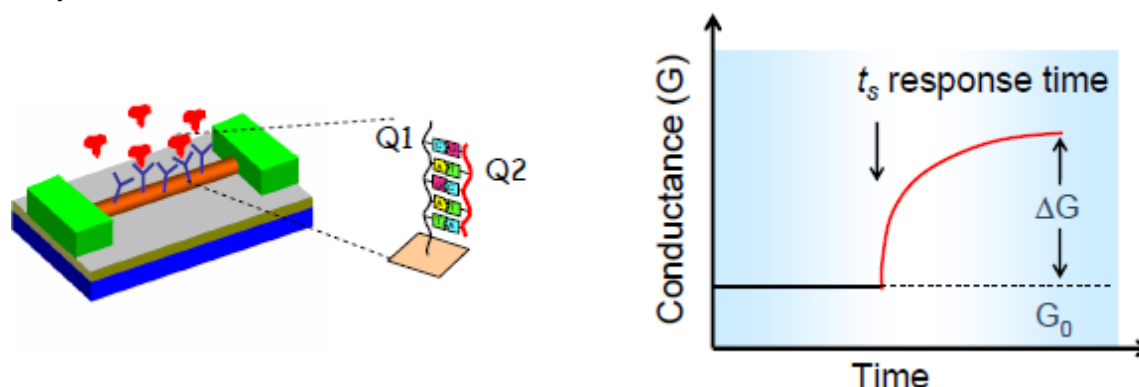


Figure 2 DNA strand and sensor response time [Landman, j. *et al* (1998)]

Thus by detecting the amount of current change, the corresponding concentration of unknown DNA sequence in a given electrolyte is detected. This is the basic principle of detection in nanobio sensor. Figure 2 shows the change in conduction of sensors due to detection of unknown sequence. There are different kinds of nanobio sensors such as Chem FET, IsFET, Nanowire, Nanosphere, Nanodots and Carbon Nano Tube [CNT]. Sensitivity is one of the major parameter that needs to be considered to select an appropriate sensor for drug delivery. Sensors consist of source and drain regions placed above a gate. Gate consists of receptors that capture the unknown molecules that diffuse across the target molecules. Figure 3 shows the two basic kinds of sensor (ISFET and Nanosensor).

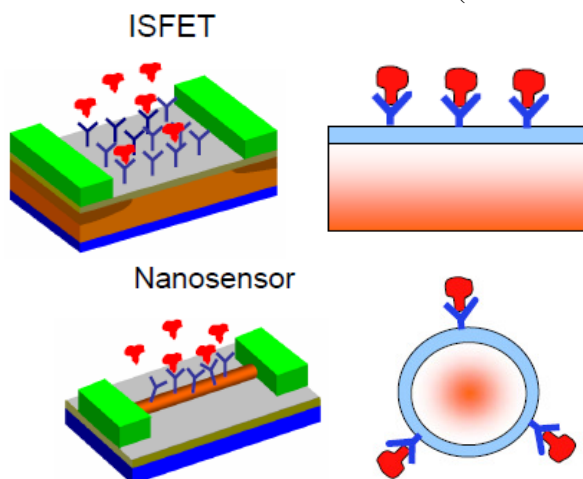


Figure 3 ISFET and Nanosensor [Landman, j. *et al* (1998)]

Current flows between source and drain, and the molecules that are bound to the sensor determine the source-drain current. The sensitivity of such sensor is found to be between molar and few micro molar (10^{-6} M) [Landman, j. *et al* (1998)]. This is a very small value. It is therefore essential that sensors should have higher sensitivity for diseases detection. To improve sensitivity of a sensor for bio applications, CNT were introduced. The sensitivity of CNT sensors compared with nanosensor was increased by several orders of magnitude (femto molar) for biomedical applications. In order to further improve sensitivity, nanodots can also be used. It is found in the available literature that the cylindrical or nanowire sensors are much better than planar sensor. The reason for this lies in the geometry of electrostatics. In a nanowire, the unknown molecules surround the gate consisting of receptor molecules as compared to a planar transistor where the receptor molecules are on top of the plane. Thus there is a higher sensitivity in nanowire. The currents in nanowires are in tens of nanometer dimension, which is very large. The cross section of nanowire sensor is shown in figure 4. The nanowire is immersed in water or pH containing material and the DNA molecules are swimming around in the electrolyte.

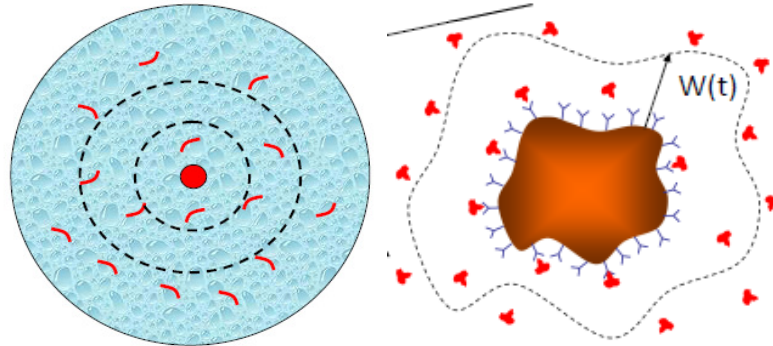


Figure 4 Diffusion region and sensor detection boundary [Landman, j. *et al* (1998)]

In order to understand or model sensor for use in drug delivery applications, it is essential to analyze the working principle of nanowire sensor and develop mathematical relationships that can be used in sensor design. A sensor consisting of numerous receptors is shown in figure 4. The unknown molecules (target) are captured by the receptors as they diffuse along the surface of receptors, only when the unknown molecule has a conjugate sequence compared to the receptor sequence. It is essential to establish the relationship between number of molecules detected, current, time involved in detection and concentration of molecules.

III. DIFFUSION-CAPTURE MODEL [LANDMAN, J. ET AL (1998)]:

There are two equations that explain the diffusion-capture activity in a nanobio sensor. The capture equation is as shown in (1):

$$\frac{dN}{dt} = k_F(N_0 - N)\rho_s - k_R N \quad [1]$$

N is the number of conjugated molecule, N_0 is the initial number of molecules (receptors, blue Y shaped). The number of conjugated is proportional to number of unconjugated molecules and is determined by $(N_0 - N)$, where k_F is reaction constant. There are possibilities of molecules that are bound to deconjugate due to chemical reaction. The second term $k_R N$ represents the number of deconjugated molecules (k_R is reverse reaction constant). Deconjugation is very weak in nanobio sensors, and hence the diffusion equation can be approximated to the equation as shown in (2):

$$\frac{dN}{dt} \approx k_F N_0 \rho_s \quad [2]$$

ρ_s is the surface concentration of the captured molecules. As the molecules present in the electrolyte diffuse across the receptors, the diffusion equation is given in (3)

$$\frac{d\rho}{dt} = D \nabla^2 \rho \quad [3]$$

D is the diffusion coefficient; ρ is the concentration of molecules. This equation defines that the molecules have to diffuse around the sensor surface before they could be captured. It is essential to establish an analytical solution for the above two equations in order to understand the sensitivity of sensors. The diffusion-capture equation [Prow, t.w. *et al* (2005), Prow, t.w. *et al* (2006), Prow, t.w. *et al* (2006)] needs to be solved to understand the behaviour of the sensor. The solution for number of molecules captured and is given in (4). This work is published in the Imanager journal [Ushaa. S.M. *et al* (2010)]

$$N(t) = \rho_0 t \left[\frac{A}{C_0} + \frac{1}{k_F N_0} \right]^{-1} \quad [4]$$

The equation (4) is used to compute the number of molecules that have been captured for a certain period of time. The capacitance C_0 is chosen based on different kind of sensor being used. Thus it can

be seen that the dimensionality of sensor influences the number of molecules captured, thereby affecting the sensitivity of the sensor. The above analysis is carried out assuming steady state analysis, i.e. the concentration of diffusion is constant within the outer boundary. In order to model the sensor behaviour in transient state, figure 5 shows a sensor at the centre, and the analyte with unknown molecules (blue). The sensor captures the molecules closer to it and as the distance increases the analyte concentration increases, and the molecules closer to the sensor are being captured (white).

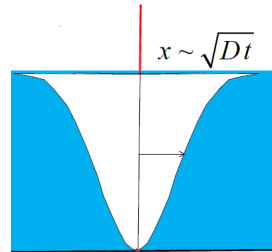


Figure 5 Diffusion changes

As the boundary of diffusion changes and is time dependent, the factor W is time dependent. The boundary surface increases with time as in figure 6.

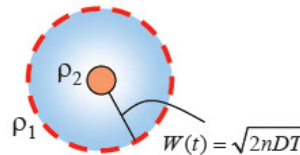


Figure 6 Variable diffusion boundaries [www.nanohub.org]

Thus the diffusion concentration is varying with time and the modified equations for $N(t)$ is given in equation in (5).

$$C_t = \frac{D}{\sqrt{2Dt}}$$

$$= \frac{2\pi D}{\log(\sqrt{4Dt} + a_0/a_0)}$$

$$= \frac{4\pi D}{a_0^{-1} - (\sqrt{6Dt} + a_0)^{-1}}$$

[5]

For different sensors as shown in figure 7, the factor W changes with the geometry.

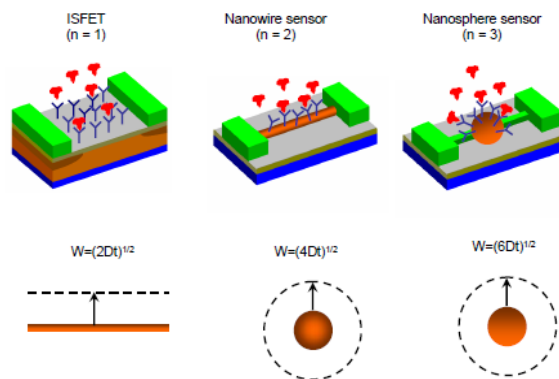


Figure 7 Different types of sensors [www.nanohub.org]

Based on the modified equation $N(t)$ given in equation (6), mathematical models for nanowire sensor is developed. The value of C_t for a transient behaviour is given in equation (5).

$$N(t) = \rho_2 t \left[\frac{A}{C_t} + \frac{1}{k_F N_0} \right]^{-1} \quad (6)$$

Choosing appropriate values for C_t , the geometries of the sensor, three different sensors can be modelled. For different sensors as shown in figure 7, the factor W changes with the geometry.

IV. EXPERIMENTAL SETUP AND SENSOR CHARACTERIZATION

Based on the mathematical models discussed, biosensor tool available in Nanohub.org is used for simulation of ISFET, nanowire and nanosphere. For a biosensor the most important parameters that are required are:

- Size of micro channel: 5mm x 0.5mm x 50um
- Flow rate of fluid in the channel: 0.15ml/h
- Concentration of antigens in fluid: $2 \cdot 10^{-15} \cdot 6 \cdot 10^{23} \approx 10^9$
- Number of antigens through channel per hour: $1.5 \times 10^{-4} \times 10^9 \sim 10^5$ (~ 42 per second)
- Total area occupied by Antibodies: 5mmx0.5mm ~ $25 \times 10^{-7} \text{m}^2$
- Area of one Si NW occupied by Antibodies (Assumption: $r \sim 10\text{nm}$, $l \sim 2\mu\text{m}$): $2\pi rl \sim 1.26 \times 10^{-15} \text{m}^2$
- Target receptor conjugation
- Type of antigen: DNA
- Ratio between total occupied area and Si NW: 2×10^9
- Mean time between one antigen reacts with one antibody on the Si NW: <3 minutes

Based on the above parameters, the parameters in the biosensor lab is developed and the models available in the sensor lab are simulated. Figure 8 shows the experimental setup using the biosensors lab for simulating three different sensors.

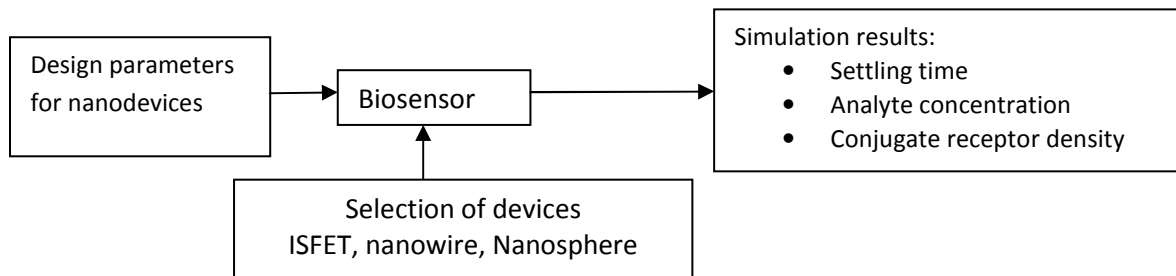


Figure 8 Experimental setup for sensor characterization

V. CHARACTERISTICS OF SILICON NANOWIRE

A silicon nanowire is developed with the following parameters, the VI characteristics of the nanowire is simulated using the biosensor tool. Sensor parameters: Diameter of silicon nanowire: 10nm, Oxide thickness: 5nm, Gate length: 50nm, Channel doping: $1 \text{e}+21 / \text{cm}^3$. Analyte concentration parameter is varied from 0.1 to 1 nmol/L, corresponding changes drain current in the nanowire sensor is determined. Figure 9 shows the graph of concentration vs. device current characteristics for nanowire sensors.

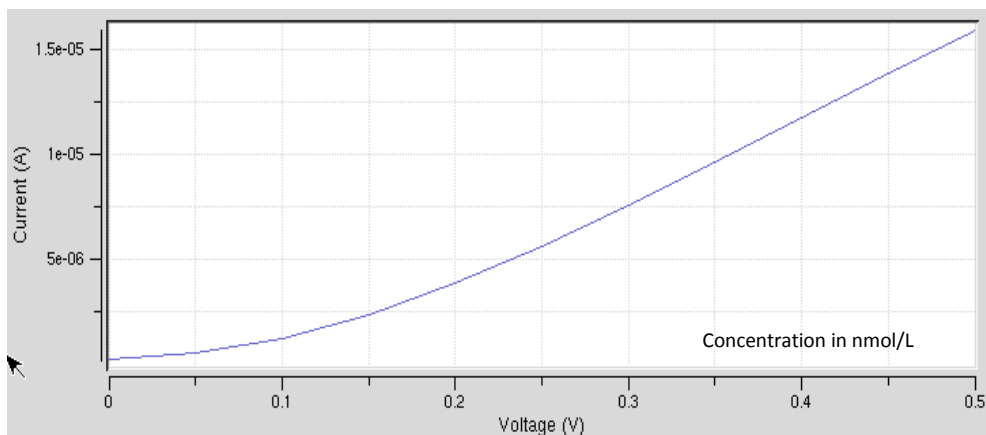


Figure 9 Concentration vs. device current

Table 1 shows the equivalents in terms of voltage samples for various sets of iterations carried out. The results have been obtained using Matlab simulations.

Table 1 Sensor equivalents

Concentration in nmol/L	Sensor current equivalents	Concentration in nmol/L	Sensor current equivalents	Concentration in nmol/L	Sensor current equivalents
0.4629	0.8332	0.3646	0.6564	0.349	0.6283
0.3706	0.667	0.2211	0.398	0.3154	0.5676
0.4616	0.8308	0.2109	0.3796	0.5427	0.9769
0.3751	0.6753	0.4324	0.7783	0.4263	0.7673
0.1138	0.2049	0.3445	0.6201	0.7965	1.4338
0.2336	0.4205	0.3558	0.6404	0.6919	1.2454
0.5667	1.02	0.1213	0.2184	0.1302	0.2343
0.6277	1.1298	0.2161	0.3891	0.124	0.2232
0.3818	0.6873	0.1137	0.2046	0.7293	1.3128
0.4559	0.8207	0.1532	0.2758	0.5636	1.0145
0.34	0.6121	0.1406	0.253	1.4003	2.5205
0.3101	0.5582	0.2606	0.469	0.6937	1.2487
0.2772	0.4989	0.235	0.423	0.4923	0.8862
0.5925	1.0665	0.116	0.2088	0.1055	0.1899
0.4978	0.8961	0.1988	0.3578	0.1297	0.2335
0.4881	0.8786	0.2067	0.372	0.9062	1.6312
0.3285	0.5913	0.0604	0.1088	0.9573	1.7231
0.3457	0.6222	0.1742	0.3136	0.387	0.6966
0.2778	0.5001	0.1478	0.2661	0.5344	0.962
0.2002	0.3604	0.1288	0.2319	0.4633	0.834
0.5852	1.0534	0.139	0.2503	0.1911	0.344
0.3123	0.5622	0.166	0.2989	0.4768	0.8582
0.583	1.0494	0.2258	0.4064	0.2374	0.4272
0.3932	0.7077	0.2193	0.3948	0.3346	0.6023
0.4084	0.7351	0.1846	0.3323	0.2624	0.4723
0.3939	0.709	0.1292	0.2326	0.5181	0.9326
0.2934	0.5281	0.261	0.4698	0.262	0.4716
0.3818	0.6872	0.2218	0.3993	0.1192	0.2146
0.5059	0.9107	0.0826	0.1487	0.0907	0.1633
0.5227	0.9408	0.2237	0.4026	0.271	0.4878
0.445	0.8011	0.1165	0.2097	0.4029	0.7252
0.307	0.5526	0.1325	0.2385	0.644	1.1592
0.1723	0.3101	0.3161	0.569	0.4642	0.8355
0.4376	0.7876	0.2097	0.3775	0.1705	0.3069
0.1768	0.3183	0.2581	0.4646	0.9265	1.6678
0.5035	0.9064	0.1799	0.3238	0.3296	0.5933
0.4297	0.7734	0.1099	0.1978	0.7458	1.3425
0.3029	0.5452	0.2529	0.4552	0.5271	0.9487
0.3945	0.7101	0.2432	0.4377	0.4593	0.8267
0.3986	0.7174	0.1736	0.3126	0.3569	0.6425

Figure 10 shows the graphical display of three different iterations of the sensor model. From the graphs shows that the variation in sensor currents is nonlinear and is also consists of noise.

A work on sensor array design for disease detection and classification is carried out and published in Inderscience Journal [Ushaa. S.M, *et al* (2011)]

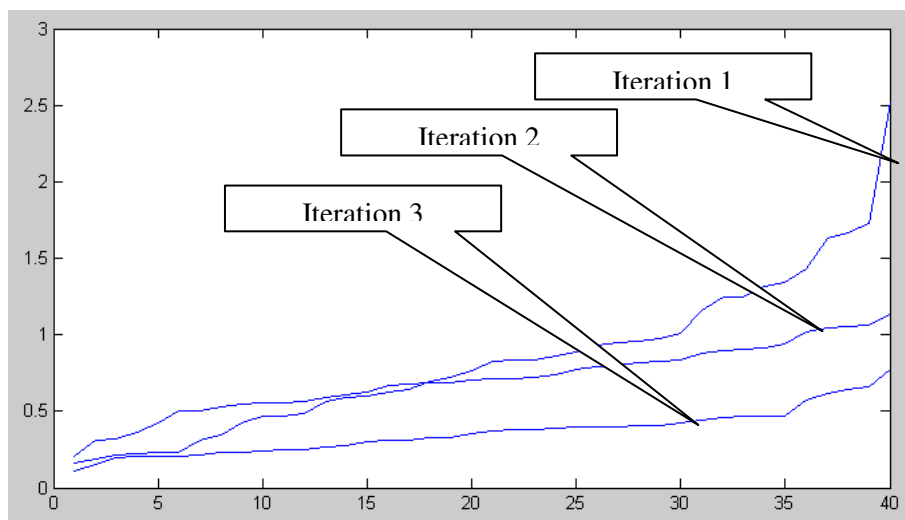


Figure 10 Concentration Vs. Sensor Currents for Three Iterations

4.2. Mathematical models for nanowire sensor:

From the results obtained using biosensor lab, a Matlab model is developed for silicon nanowire. The Matlab model is based on the results obtained in Table 1. The experimental setup developed using biosensors lab is used to identify the equivalent current values that flow in the drain of nanowire sensor with changes in analyte concentration. During the experimental setup, 135 different values of analyte concentration are set to identify the variations in drain current. The analyte concentration is varied from 0.1 to 0.5 mmol/L, corresponding drain currents are identified and recorded. The Matlab model is a look up table of these values obtained in the biosensor lab. Figure 10 shows the top level diagram of Matlab model for nanowire sensor.

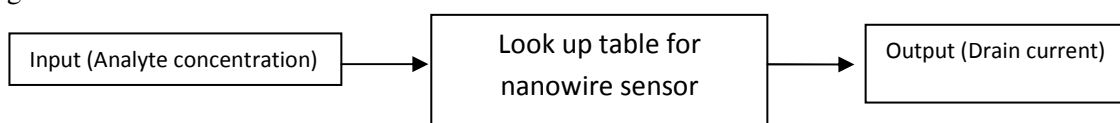


Figure.10 Matlab model developed using look up table

In order to generalize the sensor models for all possible input conditions, it is required to extend the sensor model for generic inputs. In this work, curve fitting techniques have been adopted to improve the performance characteristics of sensor models. Next section discusses the curve fitting techniques for sensor model development.

VI. CURVE FITTING TECHNIQUES

There are four different steps in curve fitting, they are 1> Data Transformation, 2> Smoothing and filtering 3> Curve Fitting and 4> Residual Analysis.

5.1 Data transformations

Changing variables through data transformations may lead to a simplified relationship between the transformed predictor variable and the transformed response. As a result, model descriptions and predictions may be simplified. Common transformations include the logarithm $\ln(y)$, and power functions such as $y^{1/2}$, y^{-1} . Using these transformations, one can linearize a nonlinear model, contract response data that spans one or more orders of magnitude, or simplify a model so that it involves fewer coefficients. In this work, as the simulation results obtained are not nonlinear throughout the span of input samples, data transformation techniques are not adopted.

5.1.1 Smoothing and filtering

There are four different filtering types they are: Moving average filter, Savitzky-Golay Filtering, Local Regression Smoothing and Smoothing Splines. Moving average filter smooths data by replacing each data point with the average of the neighboring data points defined within the span. This process is equivalent to lowpass filtering with the response of the smoothing given by the difference equation

$$y_s(i) = \frac{1}{2N+1} (y(i+N) + y(i+N-1) + \dots + y(i-N)) \quad (7)$$

where $y_s(i)$ is the smoothed value for the i th data point, N is the number of neighboring data points on either side of $y_s(i)$, and $2N+1$ is the span. Limitations of moving average filter are:

- The span must be odd.
- The data point to be smoothed must be at the center of the span.
- The span is adjusted for data points that cannot accommodate the specified number of neighbors on either side.
- The end points are not smoothed because a span cannot be defined.

Savitzky-Golay filtering can be thought of as a generalized moving average. You derive the filter coefficients by performing an unweighted linear least-squares fit using a polynomial of a given degree. For this reason, a Savitzky-Golay filter is also called a digital smoothing polynomial filter or a least-squares smoothing filter. Note that a higher degree polynomial makes it possible to achieve a high level of smoothing without attenuation of data features. The Savitzky-Golay filtering method is often used with frequency data or with spectroscopic (peak) data. For frequency data, the method is effective at preserving the high-frequency components of the signal. For spectroscopic data, the method is effective at preserving higher moments of the peak such as the line width. By comparison, the moving average filter tends to filter out a significant portion of the signal's high-frequency content, and it can only preserve the lower moments of a peak such as the centroid. However, Savitzky-Golay filtering can be less successful than a moving average filter at rejecting noise. Figure 11 shows the filter sensor current samples using Savitzky-Golay filter. From the results shown, it is found that the noise in the sensor currents are filtered and hence improves the performance of the sensor models.

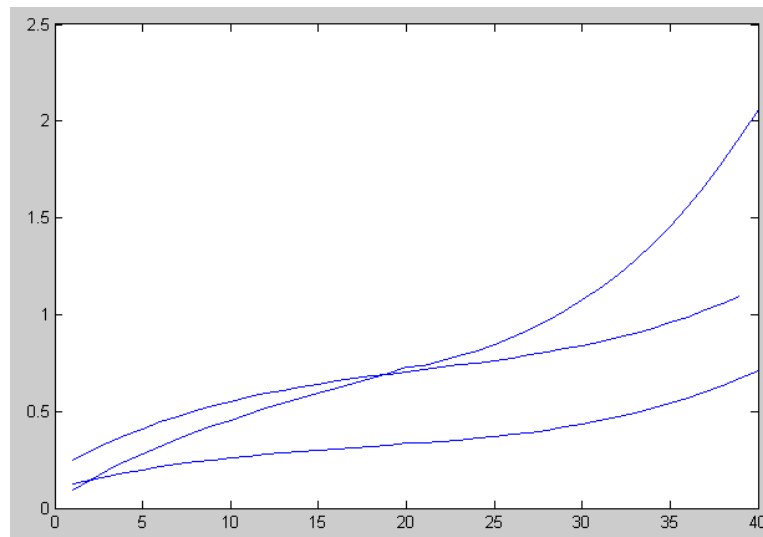


Figure 11 Savitzky-Golay Filtering of Sensor Currents

Curve fitting requires a parametric model that relates the response data to the predictor data with one or more coefficients. The result of the fitting process is an estimate of the model coefficients. To obtain the coefficient estimates, the least-squares method minimizes the summed square of residuals. The residual for the i th data point r_i is defined as the difference between the observed response value y_i and the fitted response value \hat{y}_i , and is identified as the error associated with the data and is given by:

$$r_i = y_i - \hat{y}_i$$

$$\text{residual} = \text{data} - \text{fit}$$

(8)

The summed square of residuals is given by

$$S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

where n is the number of data points included in the fit and S is the sum of squares error estimate. The supported types of least-squares fitting include:

- Linear least squares
- Weighted linear least squares
- Robust least squares
- Nonlinear least squares

A linear model is defined as an equation that is linear in the coefficients. For example, polynomials are linear but Gaussians are not. To illustrate the linear least-squares fitting process, suppose you have n data points that can be modeled by a first-degree polynomial.

$$y = p_1 x + p_2 \quad (10)$$

To solve this equation for the unknown coefficients p_1 and p_2 , you write S as a system of n simultaneous linear equations in two unknowns. If n is greater than the number of unknowns, then the system of equations is over determined.

$$S = \sum_{i=1}^n (y_i - (p_1 x_i + p_2))^2 \quad (11)$$

Because the least-squares fitting process minimizes the summed square of the residuals, the coefficients are determined by differentiating S with respect to each parameter, and setting the result equal to zero.

$$\begin{aligned} \frac{\partial S}{\partial p_1} &= -2 \sum_{i=1}^n x_i (y_i - (p_1 x_i + p_2)) = 0 \\ \frac{\partial S}{\partial p_2} &= -2 \sum_{i=1}^n (y_i - (p_1 x_i + p_2)) = 0 \end{aligned} \quad (12)$$

The estimates of the true parameters are usually represented by b . Substituting b_1 and b_2 for p_1 and p_2 , the previous equations become

$$\begin{aligned} \sum x_i (y_i - (b_1 x_i + b_2)) &= 0 \\ \sum (y_i - (b_1 x_i + b_2)) &= 0 \end{aligned} \quad (13)$$

where the summations run from $i=1$ to n . The *normal equations* are defined as

$$\begin{aligned} b_1 \sum x_i^2 + b_2 \sum x_i &= \sum x_i y_i \\ b_1 \sum x_i + n b_2 &= \sum y_i \end{aligned} \quad (14)$$

Solving for b_1

$$b_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (15)$$

Solving for b_2 using the b_1 value

$$b_2 = \frac{1}{n} (\sum y_i - b_1 \sum x_i) \quad (16)$$

As you can see, estimating the coefficients p_1 and p_2 requires only a few simple calculations. Extending this example to a higher degree polynomial is straightforward although a bit tedious. All that is required is an additional normal equation for each linear term added to the model.

In matrix form, linear models are given by the formula

$$y = X\beta + \varepsilon \quad (17)$$

where

- y is an n -by-1 vector of responses.
- β is a m -by-1 vector of coefficients.
- X is the n -by- m design matrix for the model.
- ε is an n -by-1 vector of errors.

Based on the discussions carried out in this section, least squares technique is adopted for interpolation and curve fitting of sensor model results. Sensor data captured using nanohub simulation is imported into Matlab, SG filtering is adopted to smoothen the noise from the sensor model, and is interpolated using least square curve fitting technique. Thus in this work, least square method based curve fitting technique is adopted. Based on the LS method, the residual for the i th data point r_i is defined as the difference between the observed response value y_i and the fitted response value, and is identified as the error associated with the data.

$$r_i = y_i - \hat{y}_i \quad (18)$$

residual = data – fit

From a given set of data points, the residual is computed as in equation (18), based on the residuals obtained the summed square of residuals is computed as in equation (19)

The summed square of residuals is given by

$$S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (19)$$

Knowing the residual of a curve or set of data points, an equation as given in (20), expresses the relationship between input samples x and output y .

$$y = p_1 x + p_2 \quad (20)$$

Finally the curve fitted data points are obtained as in equation [21]

$$S = \sum_{i=1}^n (y_i - (p_1 x_i + p_2))^2$$

[21]

x is the concentration of analyte solution and S is the drain current. In this work, the three iterations are expressed in terms of mathematical equations using curve fitting technique i.e $S1$, $S2$ and $S3$ are computed and finally the average of three is considered for modelling of sensor as given in equation (22)

$$\begin{aligned} S1 &= \text{sum}(y_{i1} - (p_1 x_{i1} + p_2))^2 \\ S2 &= \text{sum}(y_{i2} - (p_1 x_{i2} + p_2))^2 \\ S3 &= \text{sum}(y_{i3} - (p_1 x_{i3} + p_2))^2 \\ S_{final} &= (S1 + S2 + S3)/3 \end{aligned} \quad (22)$$

The response of nanowire sensor computed based on the model developed is used in detection of various diseases and is used in design of automated drug delivery unit. In order to validate the interpolated sensor model, a new iteration (iteration 4), is considered with change in input parameters as presented in figure 8. The sensor model parameters are varied and a new experiment is conducted to measure the variation in sensor current for different sets of molecular concentration. In figure 12, the graph (blue) is the variation in sensor currents for variation in molecular concentration. Graph (green) is the sensor current obtained based on curve fitting techniques. From the comparison of these two graphs, it is found that the variation in the actual model and the curve fitted model are similar, but there is a scaling difference. This is due to amplification factor in the mathematical equation. In order to improve the linearity between actual models and the mathematical model, SG filtering is adopted on the captured signal from the simulation results. Curve fitting techniques are adopted after filtering, thus eliminate the noise in the captured signal as well as reduce the intensity of scaling factor. The results are obtained based on SG filtering and curve fitting is shown in Figure 12. From the obtained results it is found that the curve fitted model after SG filtering matches with the actual sensor model in terms of both variation and intensity. From the results obtained it is found that the error between actual and improved sensor model is less than 0.8.

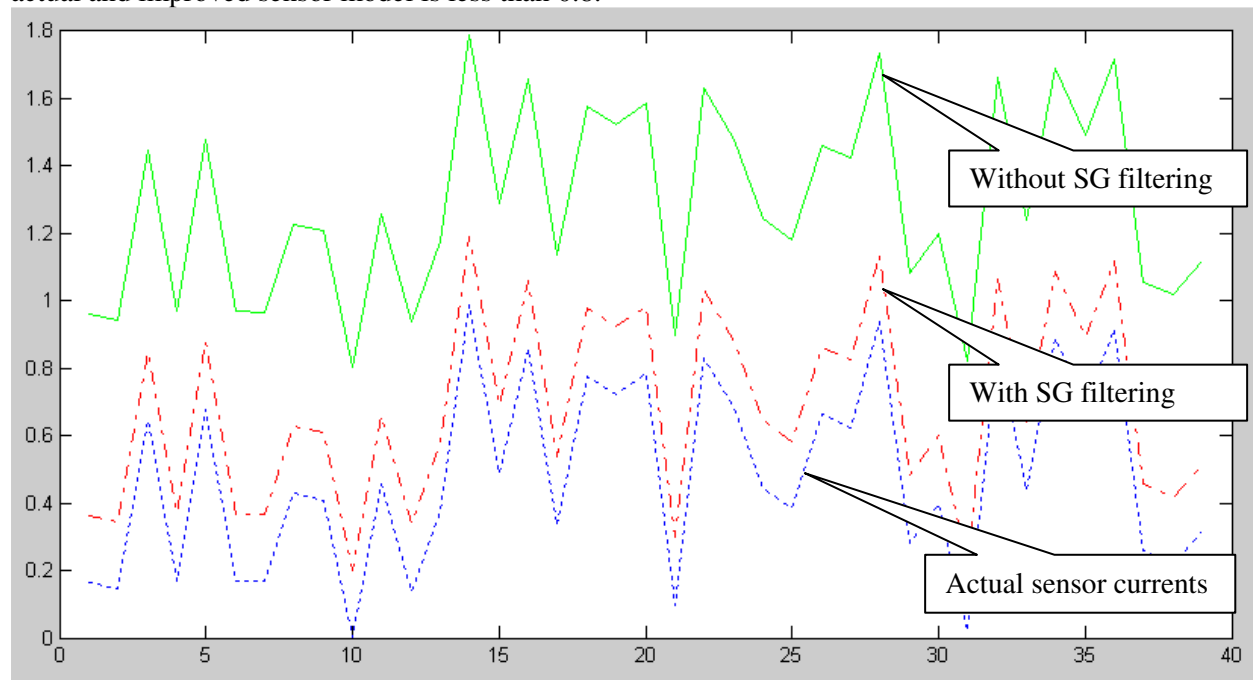


Figure 12 Performance Comparison of Sensor Models

VII. CONCLUSION

Automated disease detection and drug delivery unit that can be used for detection and monitoring of cancer is proposed, designed, modeled and implemented in this work. For the first time, a software reference model for the complete unit as a system is developed and analyzed for its functionality. An exhaustive literature review on various diseases and remedies to cure the diseases is carried out. The procedures and methods adopted by doctors to detect and diagnose patients with diseases are explored. Challenges in disease diagnosis is identified and reported in this work. It is found that with the growth in population and with changes in environmental conditions human race is prone to various diseases. In this paper, we have analyzed the mathematical models for nanowire sensors and the variation in sensor properties with geometrical parameters. Experimental setup is developed to simulate three different nanosensors (ISFET, nanowire and nanosphere). Sensitivity of nanosphere is found to be better than nanowire and ISFET, however, it is practically difficult to realize nanosphere. Thus nanowire sensor is selected for system level design (disease detection), nanowire sensor is simulated and its response to variations in analyte concentration is identified. The developed mathematical model is validated against biosensor model, the results shows that both the models have linear variations for changes in analyte concentration, but there is an error of 0.8 (maximum), between

the drain currents of biosensor model and Matlab model. This can be minimized by developing accurate results using the biosensor model for large number of analyte concentration. The mathematical model developed can be used to model different sensors using Matlab. The sensors can be interfaced with signal conditioning circuits and control unit for automated disease detection and drug delivery.

ACKNOWLEDGEMENT

The authors would like to acknowledge Nanohub.org for providing permission to access the biosensor labs simulation tools and to carry out the experiments. The authors also acknowledge the support and guidance provided by Cyril Prasanna Raj P., MSR School of Advanced Studies, Bangalore. His inputs and timely guidance has helped us to carry out the experimental analysis.

REFERENCES

- [1]. Beckett, m.; cazares, l.; vlahou, a.; schellhammer, p.; wright, g.(1995) Prostate-specific membrane antigen levels in sera from healthy men and patients with benign prostate hyperplasia or prostate cancer. *Clin. Cancer res.* 1995, 5, 4034-4040.
- [2]. Catalona, w. (1996)Clinical utility of measurements of free and total prostate-specific antigen (psa): a review. *Prostate*, 7, 64-69.
- [3]. Dandachi, n.; dietze, o.; hauser-kronberger, c.(2002) Chromogenic in situ hybridization: a novel approach to a practical and sensitive method for the detection of her2 oncogene in archival human breast carcinoma. *Lab. Invest.* 2002, 82, 1007–1014.
- [4]. Henderson, c.; patek, a.(1998) The relationship between prognostic and predictive factors in the management of breast cancer. *Breast cancer res. Treat.* 1998, 52, 261-288
- [5]. Ludwig, j.; weinstein, j. (2005)Biomarkers in cancer staging, prognosis and treatment selection. *Nature rev. Cancer*, 5, 845-856.
- [6]. Landman, j.; chang, y.; kavalier, e.; droller, m.; liu, b.(1998) Sensitivity and specificity of nmp-22, telomerase, and bta in the detection of human bladder cancer. *Urology* 1998, 52, 398-402.
- [7]. Molina, r.; auge, j.; escudero, j.; marrades, r.; vinolas, n.; carcereeny, e.; ramirez, j.; filella, x. Mucins ca(2008) 125, ca 19.9, ca 15.3 and tag-72.3 as tumor markers in patients with lung cancer: comparison with cyfra 21-1, cea, scc and nse. *Tumor biol.* 2008, 29, 371-380.
- [8]. Prow, t.w., rose, w.a., wang, n., reece, l.m., lvov, y., leary, j.f (2005). "biosensor-controlled gene therapy/drug delivery with nanoparticles for nanomedicine" *proc. Of spie* 5692: 199 –208, 2005.
- [9]. Prow, t.w., smith, j.n., grebe, r., salazar, j.h., wang, n., kotov, n., luty, g., leary, j.f.(2006) "construction, gene delivery, and expression of dna tethered nanoparticles" *molecular vision* 12: 606-615, 2006.
- [10]. Prow, t.w., grebe, r., merges, c., smith, j.n., mcleod, d.s., leary, j.f., gerard a. Luty, g.a.(2006) "novel therapeutic gene regulation by genetic biosensor tethered to magnetic nanoparticles for the detection and treatment of retinopathy of prematurity" *molecular vision* 12: 616-625, 2006.6.
- [11]. Seeman (2005) from genes to machines: dna nanomechanical devices. *Trends in biochemical sciences* 30
- [12]. Ushaa eswaran, m. Madhavalatha, madhusudhana rao ganji.(2006) A proposal for design of hybrid multi-functional biochip using bio-intelligent expert system' the 2006 international conference on artificial intelligence (icai'06) june 26-29, , las vegas, usa
- [13]. Ushaa eswaran, madhusudhan rao ganji , jayashankarts, lekha suresh,(2006) 'design of implantable vlsi bio-chip', national conference on advanced communication technologies[act-06],8-9 december 2006.
- [14]. Ushaa eswaran, m.madhavalatha, , (2009)"disease detection using pattern recognition techniques" national conference on 'emerging trends in information communication technology' (etict-08) held in gitam university, india during 19th & 20th december and is published in gitam journal of information communication technology of vol-2 jan - july 2009 number - 1 (issn 0974-4622) pp 34-37
- [15]. Ushaa eswaran, madhusudhana rao, m.s.thakur (2004) 'microprocessor based biosensors for determination of toxins and pathogens in restricted areas of human intervention' ic ai'04 ieee sponsored international conference on artificial intelligence held in las vegas, nevada, usa during 21 – 24 june 2004.
- [16]. Ushaa. S.M.; madhavalatha.m; madhusudhana rao ganji.(2010) - "development and validation of matlab models for nanowire sensors for disease detection" *i-manager's journal on future engineering & technology*, vol. 6 1 no. 2 1
- [17]. Ushaa. S.M.; Madhavalatha.M; Madhusudhana Rao Ganji (2011)-"Design and Analysis of

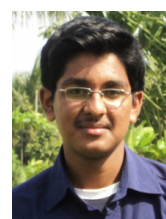
Nanowire Sensor Array for Prostate Cancer Detection" (Submission code: IJNBM-20274) for the International Journal of Nano and Biomaterials (IJNBM). *Int. J. Nano and Biomaterials*, Vol. 3, No. 3. [18]. www.nanohub.org

Author's Profile

S. M. Usha obtained her Bachelors Degree in Electronics and Instrumentation Engineering from the University of Annamalai in the year 1989 and her post graduation degree in Electronic Instrumentation Engineering from the University of Andhra, Vishakhapatnam in the year 2003. She is presently pursuing her Ph.D. from the Jawaharlal Nehru Technological University Hyderabad. She has obtained with distinction the International certificate for teachers and trainers being awarded by Cambridge University.



Vivek Eswaran is presently pursuing his Bachelors degree in Electronics and Communication Engineering from Anna University. A Pre Final year student of RMK Engineering College, Vivek is an academically brilliant student having consistently securing distinction in all his previously held examinations. Vivek has always exhibited interest in topics that have potential for revolutionizing sciences. His electronic background has helped in conceiving and exploring mathematical models and simulation studies.



MISSING BOUNDARY DATA RECONSTRUCTION BY AN ALTERNATING ITERATIVE METHOD

Chakir Tajani and Jaafar Abouchabaka
LARIT, Ibn Tofail University, Kenitra , Morocco

ABSTRACT

In this work, we consider the Cauchy problem for Laplace equation which is to complete the missing data on a part of the boundary from the over specified conditions on the other part. Among the existing method to solve this ill-posed problem, we find the iterative algorithm (KMF standard algorithm) proposed by Kozlov, Mazya and Fomin. This algorithm requires a large number of iterations to complete convergence. Therefore, we propose an alternating iterative algorithm (KMF developed algorithm) to ameliorate the accuracy of the solution and reduce the number of iterations needed to achieve convergence. This algorithm is inspired by the KMF standard algorithm and the numerical tests performed to study the influence of data problem on the convergence rate. The implementation of this algorithm is performed by the finite element method using the software Freefem. The numerical tests developed show the effectiveness of the proposed algorithm since it allows to have more accurate results as well as reducing the number of iterations needed for convergence.

KEYWORDS: Inverse problem, Laplace equation, iterative method, Freefem.

I. INTRODUCTION

We consider an inverse problem for Laplace equation called data completion problem, which aims at recovering missing conditions on some inaccessible part of the boundary (which cannot be evaluated due to the physical difficulties or inaccessibility geometric) from the over specified boundary data on the remaining part of the boundary.

This kind of problem arises in many industrial, engineering or biomedical applications under various forms: identification of boundary conditions or physical parameters entering into its formulation, expansion of measured surface fields inside a body from partial boundary measurements. This last application can be the first step in general parameter identification problems where only partial boundary data are available. The more common problem, borrowed from thermostatic, consists in recovering the temperature in a given domain when its distribution and the heat flux are known over the accessible region of the boundary [1],[2].

This problem is ill-posed in the sense of Hadamard [3], since the existence, uniqueness and stability of the solution are not always assured [4],[5],[6]. Solving this problem by direct method is very difficult and leads to unstable solutions because a small perturbation of data provides a large difference between the solution obtained by disturbed data and that obtained by undisturbed data. Many performing numerical methods have been developed to overcome the ill-posed nature of this kind of problem. There are two types of methods: probabilistic [7] and deterministic [8].

Among the deterministic method we find, the lost ancient of them is the one, based on optimization tools, introduced by Kohn and Vogelius [9], we mention the method of Quasi-reversibility introduced by Lates since 1960 [10] , and recently used by [11],[12], Thikhonv method [13] and the iterative method.

The group of iterative method has the advantage to allow any physical constraints to be easily taken into account directly in the scheme of the iterative algorithm and simplicity of the implementation schemes. One possible disadvantage of this kind of method is the large number of iterations that may be required in order to achieve convergence.

The (KMF) algorithm proposed by Kozlov, Mazya and Fomin [14] is a very interesting iterative method developed since 1991 and has been the subject of several theoretical and numerical studies [15],[16],[17]. This method consists to solving alternately two well-posed problems in order to determine a convergent sequence to missing data on the inaccessible part of the boundary. However, its only inconvenience is the large number of iterations needed to achieve the convergence. To overcome this problem, relaxation methods have been developed by introducing a relaxation parameter in the Dirichlet condition or Neumann obtained after solving the two problems posed of the algorithm [18]. In addition, studies were made on the choice of the optimum parameter to ameliorate the rate of convergence of the algorithm [19].

In this paper, we propose an alternative iterative method based on the (KMF) method and the numerical tests performed on the influence of data problem on the rate of convergence [20], showing that the number of iterations decreases if the inaccessible part of the boundary is more smaller and two conditions are not needed on the whole accessible part to complete the remaining data. This algorithm that we call “KMF developed Algorithm” reduces considerably the number of iterations required to achieve convergence with greater accuracy.

After a presentation of the Cauchy problem for Laplace equation and a description of the mathematical algorithm proposed by Kozlov, Mazya and Fomin, the KMF developed algorithm is presented, see Section 2. Furthermore, in Section 3, various types of convergence, accuracy criteria and stopping criterion for the algorithm are considered. Many experiments are presented, including the comparison between the KMF algorithm standard and the KMF developed algorithm. The numerical results reveal several features of the alternative iterative numerical implementation that is proposed.

II. MATHEMATICAL FORMULATION

Consider an isotropic homogeneous medium consisting of an open bounded domain Ω with a piecewise smooth boundary where $\Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, $mes(\Gamma_0) \neq 0$ and $mes(\Gamma_2) \neq 0$, $mes(\Gamma_2) \geq mes(\Gamma_0)$.

In the direct problem formulation, if the temperature T and the heat flux $\partial_n T$ are known, then the temperature distribution in the domain can be calculated. However, many experimental impediments may arise in measuring or enforcing a complete boundary specification over the whole boundary Γ .

The problem is to construct a harmonic function u solution of the Cauchy problem for the Laplace equation:

$$\begin{cases} \Delta T = 0 & \text{on } \Omega \\ T_{/\Gamma_1} = f, \quad \partial_n T_{/\Gamma_2} = h, \quad T_{/\Gamma_2} = g, \quad T_{/\Gamma_3} = k \end{cases} \quad (1)$$

where f and g are prescribed functions and n is the outward normal vector to the boundary Γ . In the above formulation of the boundary conditions (2), it can be seen that the boundary Γ_2 is over specified by prescribing both the temperature g and the heat flux h ; whilst the boundary Γ_0 is underspecified since both the temperature $T_{/\Gamma_0}$ and the heat flux $\partial_n T_{/\Gamma_0}$ are unknown and have to be determined.

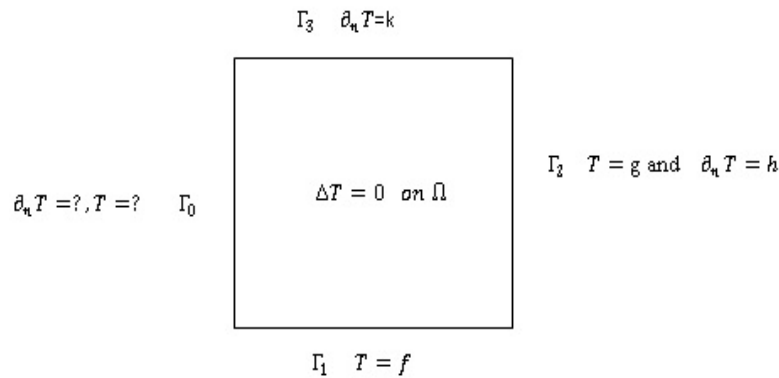


Figure 1. Schematic diagram showing the prescription of the boundary conditions (2)

Given the ill-posed aspect of this problem, because its solution if it exists is unstable; this problem cannot be solved by direct methods. However, the problem has a unique solution on $H^{(1)}(\Omega)$ when the problem data are compatible; but this solution unstable when the data on Γ_2 are disturbed. Therefore, we use a convergent iterative algorithm first proposed by Kozlov, Mazya and Fomin since 1991 which is described in Section 2.1.

At this stage, it is also worth mentioning that the analysis of this study related to the Laplace Eq. (1) can be extended to nonlinear heat conduction governed by $\nabla \cdot (k(T) \nabla T) = 0$ where $k(T)$ is the thermal conductivity [21].

2.1. Description of the KMF standard algorithm

Consider the Cauchy problem (1),(2) with $f \in H^{\frac{1}{2}}(\Gamma_1)$, $g \in H^{\frac{1}{2}}(\Gamma_2)$, $h \in H^{\frac{1}{2}}(\Gamma_2)$, $k \in H^{\frac{1}{2}}(\Gamma_3)$. The iterative algorithm investigated is based on reducing this ill-posed problem to a sequence of mixed well-posed boundary value problems and consists of the following steps:

Step 1. Specify an initial guess $u_0 \in H^{\frac{1}{2}}(\Gamma_0)$ for the temperature T on the boundary Γ_0 .

Step 2. Solve the following mixed well-posed boundary value problem:

$$\begin{cases} \Delta T^{(0)} = 0 & \text{on } \Omega \\ T^{(0)}|_{\Gamma_0} = u_0, \quad \partial_n T^{(0)}|_{\Gamma_2} = h, \quad T^{(0)}|_{\Gamma_1} = f, \quad T^{(0)}|_{\Gamma_3} = k \end{cases} \quad (3)$$

$$\text{To obtain } v_0 = \partial_n T^{(0)}|_{\Gamma_0} \quad (4)$$

Step 3. For $n \geq 1$, solve alternatively the two mixed well-posed boundary value problems:

$$\begin{cases} \Delta T^{(2n-1)} = 0 & \text{on } \Omega \\ \partial_n T^{(2n-1)}|_{\Gamma_0} = v_{n-1}, \quad T^{(2n-1)}|_{\Gamma_2} = g, \quad T^{(2n-1)}|_{\Gamma_1} = f, \quad \partial_n T^{(2n-1)}|_{\Gamma_3} = k \end{cases} \quad (5)$$

$$\text{To obtain } u_n = T^{(2n-1)}|_{\Gamma_0} \quad (6)$$

and

$$\begin{cases} \Delta T^{(2n)} = 0 & \text{on } \Omega \\ T^{(2n)}|_{\Gamma_0} = u_n, \quad \partial_n T^{(2n)}|_{\Gamma_2} = h, \quad T^{(2n)}|_{\Gamma_1} = f, \quad \partial_n T^{(2n)}|_{\Gamma_3} = k \end{cases} \quad (7)$$

$$\text{To obtain } v_n = \partial_n T^{(2n)}|_{\Gamma_0} \quad (8)$$

Step 4. Repeat step.3 from $n \geq 1$ until a prescribed stopping criterion is satisfied

2.2. Remarks

- Kozlov, Mazya et Fomin showed that for $f \in H^{\frac{1}{2}}(\Gamma_1)$, $g \in H^{\frac{1}{2}}(\Gamma_2)$, $k \in H^{-1/2}(\Gamma_3)$

and $h \in H^{-1/2}(\Gamma_2)$, the KMF standard algorithm that consists of the steps described above produces two convergent sequences both in $H^1(\Omega)$ to the solution T of the Cauchy problem (1), and that for any initial guess $u_0 \in H^{\frac{1}{2}}(\Gamma_0)$.

- The same conclusion can be obtained if at the step 1, one considers a given initial guess the form $v_0 \in H^{-\frac{1}{2}}(\Gamma_0)$ instead of the initial guess $u_0 \in H^{\frac{1}{2}}(\Gamma_0)$, and modifies accordingly the steps (2) and (3) of the algorithm.
- The algorithm would not converge if in the step 3 the mixed problems were replaced by Dirichlet or Neumann problems.
- The algorithm was found to produce an accurate and stable numerical solution for the Cauchy problem. However, one possible disadvantage of the method is the large number of iterations necessary to achieve convergence if the initial guess is far from the exact solution.

2.3. Description of the KMF developed algorithm

In this study we develop a new algorithm called “KMF developed Algorithm” in order to improve the rate of convergence of the iterative algorithm described.

The numerical results obtained by studying the influence of data problems, particularly; the relationship between the measure of the inaccessible part of the boundary and the rate of convergence have shown that:

- The convergence is always guaranteed and it is very fast if the measure of the part of the boundary to be completed is small.
- The convergence requires much more iterations if the inaccessible part is greater.

The main idea of the algorithm alternative proposed is based on the use of the previous results and the KMF standard algorithm by completing the missing data in alternative way to the two sub-parts of the inaccessible boundary. The inaccessible part is subdivided in two parts, and the KMF standard algorithm is used to complete the data in the first part, then to complete the data in the second part in an alternative way. For this, we consider $\Gamma_0 = \Gamma_{0,1} \cup \Gamma_{0,2}$ such that $\Gamma_{0,1} \cap \Gamma_{0,2} = \emptyset$ and

$$mes(\Gamma_{0,1}) = mes(\Gamma_{0,2})$$

The algorithm consists of the following steps:

Step 1. Specify an initial guess $u_0 \in H^{\frac{1}{2}}(\Gamma_0)$ for the temperature T on the boundary Γ_0 .

Step 2. Solve the following mixed well-posed boundary value problem:

$$\begin{cases} \Delta T^{(0)} = 0 & \text{on } \Omega \\ T^{(0)}|_{\Gamma_0} = u_0, \quad \partial_n T^{(0)}|_{\Gamma_2} = h, \quad T^{(0)}|_{\Gamma_1} = f, \quad T^{(0)}|_{\Gamma_3} = k \end{cases} \quad (9)$$

$$\text{To obtain } v_{0,1} = \partial_n T^{(0)}|_{\Gamma_{0,1}} \text{ and } v_{0,2} = \partial_n T^{(0)}|_{\Gamma_{0,2}} \quad (10)$$

Step 3. For $n \geq 1$, solve the two mixed well-posed boundary value problems:

$$\begin{cases} \Delta T^{(2n-1)} = 0 & \text{on } \Omega \\ \partial_n T^{(2n-1)}|_{\Gamma_{0,1}} = v_{1,n-1}, \quad \partial_n T^{(2n-1)}|_{\Gamma_{0,2}} = v_{2,n-1}, \quad T^{(2n-1)}|_{\Gamma_1} = f, \quad T^{(2n-1)}|_{\Gamma_2} = f, \quad \partial_n T^{(2n-1)}|_{\Gamma_3} = k \end{cases} \quad (11)$$

$$\text{To obtain } u_{1,n} = T^{(2n-1)}|_{\Gamma_{0,1}} \quad (12)$$

$$\begin{cases} \Delta K^{(2n-1)} = 0 & \text{on } \Omega \\ \partial_n K^{(2n-1)}|_{\Gamma_{0,1}} = u_{1,n}, \quad \partial_n K^{(2n-1)}|_{\Gamma_{0,2}} = v_{2,n-1}, \quad K^{(2n-1)}|_{\Gamma_1} = f, \quad K^{(2n-1)}|_{\Gamma_2} = g, \quad \partial_n K^{(2n-1)}|_{\Gamma_3} = k \end{cases} \quad (13)$$

$$\text{To obtain } u_{2,n} = K^{(2n-1)}|_{\Gamma_{0,2}} \quad (14)$$

Next, we solve the two mixed well-posed boundary value problem

$$\begin{cases} \Delta T^{(2n)} = 0 & \text{on } \Omega \\ T^{(2n)}_{/\Gamma_{0,1}} = u_{1,n}, \quad T^{(2n)}_{/\Gamma_{0,2}} = u_{2,n}, \quad T^{(2n)}_{/\Gamma_1} = f, \quad \partial_n T^{(2n)}_{/\Gamma_2} = h, \quad \partial_n T^{(2n)}_{/\Gamma_3} = k \end{cases} \quad (15)$$

$$\text{To obtain } v_{1,n} = \partial_n T^{(2n)}_{/\Gamma_{0,1}} \quad (16)$$

$$\begin{cases} \Delta K^{(2n)} = 0 & \text{on } \Omega \\ K^{(2n)}_{/\Gamma_{0,1}} = v_{1,n}, \quad K^{(2n)}_{/\Gamma_{0,2}} = u_{2,n}, \quad K^{(2n)}_{/\Gamma_1} = f, \quad \partial_n K^{(2n)}_{/\Gamma_2} = h, \quad \partial_n K^{(2n)}_{/\Gamma_3} = k \end{cases} \quad (17)$$

$$\text{To obtain } v_{2,n} = \partial_n K^{(2n)}_{/\Gamma_{0,2}} \quad (18)$$

Step 4. Repeat step.3 from $n \geq 1$ until a prescribed stopping criterion is satisfied.

If we consider every iteration to consist of solving the four mixed well-posed problems from the Step 2 of the algorithm, then for every $n \geq 1$ the following approximations are obtained at the iteration number n :

$T^{(2n-1)}$ and $K^{(2n-1)}$ for the temperature inside the solution domain Ω .

$T^{(2n)}$ and $K^{(2n)}$ for the temperature inside the solution domain Ω .

$u_{1,n}$ for the temperature on the $\Gamma_{0,1}$

$u_{2,n}$ for the temperature on the $\Gamma_{0,2}$

$v_{1,n}$ for the heat flux through the boundary $\Gamma_{0,1}$

$v_{2,n}$ for the heat flux through the boundary $\Gamma_{0,2}$

It should be noted that:

- ✓ The KMF developed algorithm can be seen as two parallel problems of KMF standard algorithm. These two problems are initialized with the same initial data. Each problem allows to obtain approximation in each subpart $\Gamma_{0,i}$ ou $i = 1, 2$ (for the approximation in $\Gamma_{0,1}$ the two well-posed problems (11) and (15), for the approximation in $\Gamma_{0,2}$ the two well-posed problems (13) and (17)).
- ✓ Each solved problem allows an approximation in one of the inaccessible subparts that can be introduced in the other well-posed problems.
- ✓ The missing Dirichlet condition in the part Γ_0 can be obtained from the problem (13) since the condition $u_{1,n} = T^{(2n-1)}_{/\Gamma_{0,1}}$ obtained from (11) is introduced in (13) which also provides $u_{2,n} = K^{(2n-1)}_{/\Gamma_{0,2}}$.
- ✓ The missing Neumann condition in the part Γ_0 can be obtained from the problem (17), since the condition $v_{1,n} = \partial_n T^{(2n)}_{/\Gamma_{0,1}}$ obtained from (15) is introduced in (17) which also provides $v_{2,n} = \partial_n K^{(2n)}_{/\Gamma_{0,2}}$.

III. NUMERICAL RESULT AND DISCUSSION

In this section, as a typical example, we consider the example with non-smooth boundary, such as a square $\Omega = (0, L) \times (0, L)$ where $L = 1$, namely, the analytical harmonic temperature to be retrieved is given by:

$$T(x, y) = \cos(x) \cosh(y) + \sin(x) \sinh(y). \quad (19)$$

The same example was used by Lesnic et al. [3] and was chosen to ensure that all boundary data is not zero.

For the implementation of the iterative algorithm we use the software FreeFem, and solve the well-posed problems in the algorithm by the finite element method. In this example, we use a finite element method with continuous piecewise linear polynomials to provide simultaneously the unspecified boundary temperatures and the heat fluxes.

The convergence of the algorithm may be investigated by evaluating at every iteration the error

$$e_u = \|u_n - T_{ex/\Gamma_0}\|_{L^2(\Gamma_0)} \quad \text{and} \quad e_v = \|v_n - \partial_n T_{ex/\Gamma_0}\|_{L^2(\Gamma_0)} \quad (20)$$

where u_n and v_n is the approximation obtained for the function on the boundary Γ_0 after n iterations and u_{ex} is the exact solution of the problem (1) and (2). However, in practical applications the error e_u cannot be evaluated since the analytical solution is not known and therefore the error

$$E = \|u_{n+1} - u_n\|_{L^2(\Gamma_0)} \leq \varepsilon \quad (21)$$

where ε is a small prescribed positive quantity has to be used. Alternative stopping criteria can be found in [11].

The under-specified boundary was taken to be $\Gamma_0 = \{0\} \times (0, L)$, while the over-specified boundary is $\Gamma_2 = \{L\} \times (0, L)$. The known boundaries are given by $\Gamma_3 = (0, L) \times \{L\}$ and $\Gamma_1 = (0, L) \times \{0\}$.

The known conditions are:

$$T_{/\Gamma_1} = \cos(x) = h, \quad T_{/\Gamma_2} = \cos(L) \cosh(y) + \sin(L) \sinh(y) = f$$

$$\partial_n T_{/\Gamma_2} = -\sin(L) \cosh(y) + \cos(L) \sinh(y) = g$$

$$\partial_n T_{/\Gamma_3} = \cos(x) \sinh(L) + \sin(x) \cosh(L) = k$$

And the unknown conditions in the inaccessible part and that can be determined

$$f_0 = T_{/\Gamma_0} = \cosh(y) \quad \text{and} \quad g_0 = \partial_n T_{/\Gamma_0} = -\sinh(y)$$

An arbitrary function u_0 may be specified as an initial guess for the initial temperature values, but in order to improve the rate of convergence of the iterative procedure, we have chosen a function which ensures the continuity of the temperature at the endpoints of the space interval and which is also linear with respect to the spatial coordinate x . Then, in this example, we consider the initial guess given by

$$u_0 = 1 + y(-L + \sinh(L)) + \frac{y^2}{2}, \quad y \in [0, 1] \quad (22)$$

that is not too close to the exact solution.

To implement the KMF developed algorithm, we divided the inaccessible part Γ_0 in two disjointed parts $\Gamma_{0,1}$ and $\Gamma_{0,2}$, such that $\Gamma_0 = \Gamma_{0,1} \cup \Gamma_{0,2}$ with $\Gamma_{0,1} \cap \Gamma_{0,2} = \emptyset$ and $mes(\Gamma_{0,1}) = mes(\Gamma_{0,2})$

We take $\Gamma_{0,1} = \{0\} \times (0, l)$ and $\Gamma_{0,2} = \{0\} \times (l, L)$ with $l = \frac{1}{2}$

The figure 2. and figure 3. Present a comparison between the numerical results e_u and e_v obtained with the KMF standard algorithm and the KMF developed algorithm.

It can be seen that the algorithm proposed decreases considerably the number of iteration necessary to achieve the convergence that can be reduced, and present a more accurate approximations for both Dirichlet and Neumann missing data.

For the KMF standard algorithm, the error e_u after the final iteration 43 is $3.5 \cdot 10^{-3}$. However, with the algorithm proposed it is reduced to $8,7 \cdot 10^{-4}$ after 35 iterations.

For the KMF standard algorithm the error e_v after the final iteration 42 is $16.1 \cdot 10^{-3}$. However, with the KMF developed algorithm it is reduced to $7.4 \cdot 10^{-3}$ after 24 iterations.

Then the new algorithm is more accurate to approximate the Dirichlet and Neumann conditions in the inaccessible part.

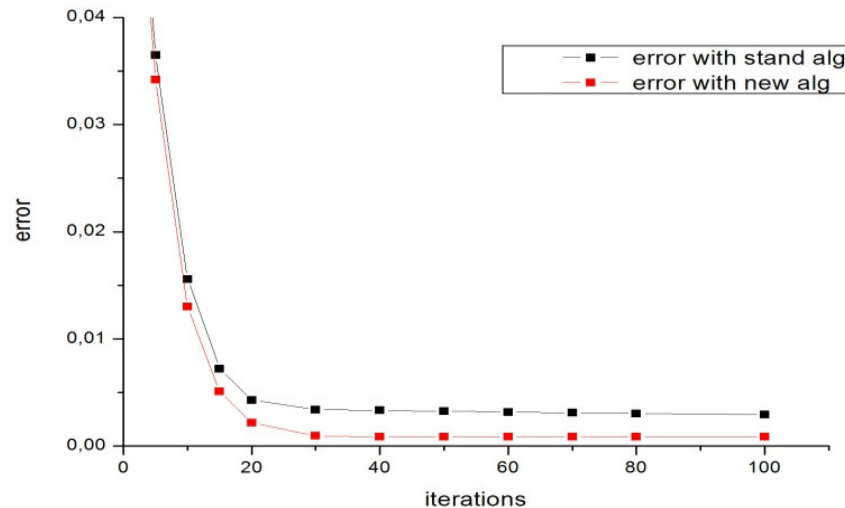


Figure 2. The error e_u obtained with the KMF standard algorithm and the KMF developed algorithm functions of the number of iterations.

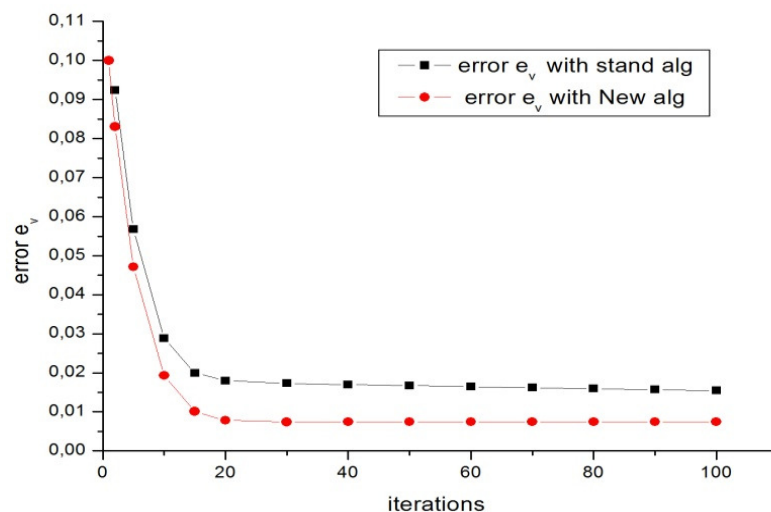


Figure 3. The error e_v obtained with the KMF standard algorithm and with the KMF developed algorithm functions of the number of iterations.

Figure 4. shows the numerical results obtained in approximating the temperature in the part of the boundary Γ_0 , indicating that from a choice of an initial data, we obtain satisfying results for both algorithms. However, the KMF developed algorithm requires less iteration to achieve more accurate convergence.

Figure 5. shows the numerical results obtained in approximating the heat flux in the part of the boundary Γ_0 by the KMF developed algorithm and the KMF standard algorithm in comparison with the analytical solution $g_0 = \partial_n T|_{\Gamma_0}$. The approximating obtained are considered more than reasonable, bearing in mind that we have solved a difficult ill-posed problem in a non-smooth domain which the improving of the accuracy and the decrease in the rate of convergence might have been manifested.

Such implementation of the algorithm allows us to notice that after the number of iterations is sufficiently increased, the error become small; this shows that the numerical solution is accurate and consistent with the number of iterations.

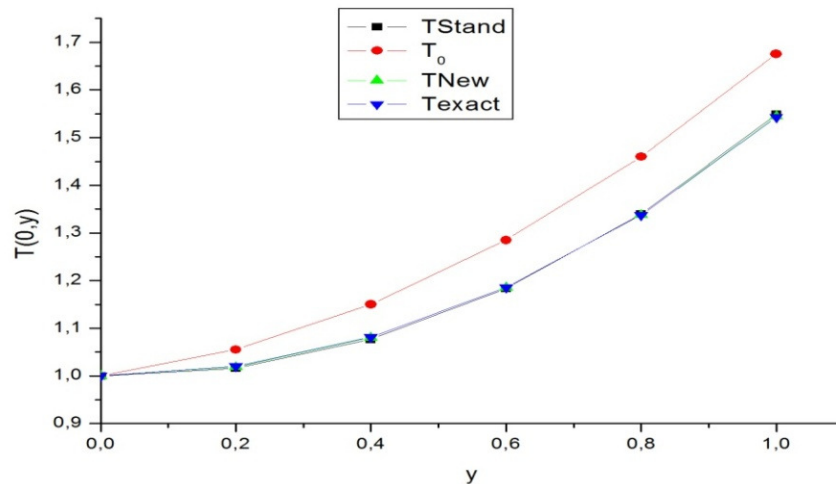


Figure 4. The numerical results for the temperature on the boundary Γ_0 obtained with KMF developed algorithm (TNew) in comparison with the analytical solution (Texact), the initial guess (u_0) and the solution with KMF algorithm (Tstand)

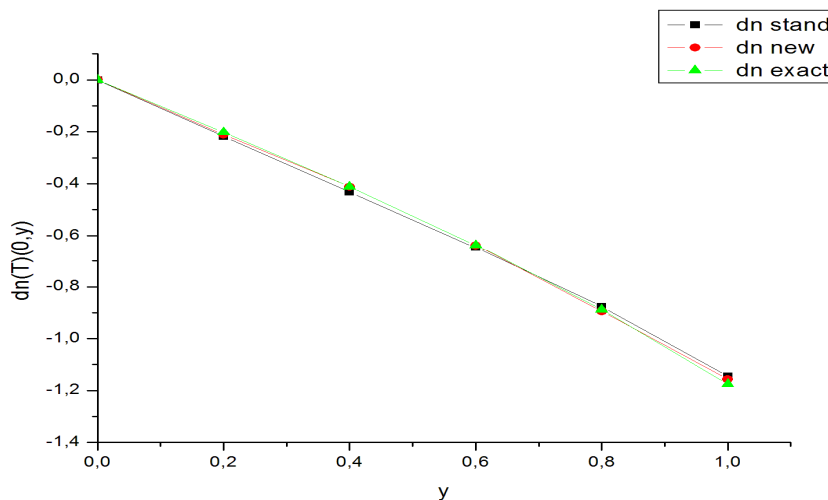


Figure 5. The numerical results obtained using the KMF developed algorithm (dn new) and the result obtained using the KMF standard (dn stand) algorithm in comparison with the analytical solution $g_0 = \partial_n T|_{\Gamma_0}$ (dn exact)

IV. CONCLUSIONS

In this work, an alternative iterative algorithm is proposed for solving a Cauchy problem for the Laplace equation. Comparison of numerical results with those obtained by the standard algorithm KMF show that the proposed algorithm significantly reduces the number of iterations needed to achieve the convergence and produces more accurate results. In addition, it can be concluded that the proposed algorithm is very efficient to reduce the rate of convergence. When perturbations are introduced into the given data problem the results are stable. Overall, it can be concluded that the alternating iterative algorithm (called KMF developed algorithm) proposed produces a convergent, stable and accurate numerical solution.

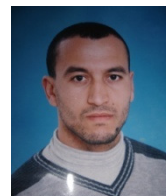
In this paper, the part of the boundary is divided in two parts. Work is in progress for implementing this algorithm by dividing the inaccessible part in more than two parts, and implementing the new algorithm in the case of the Cauchy problem for Helmholtz.

REFERENCES

- [1] D. Lesnic, L. Elliott & D.B. Ingham, (1997) “An iterative boundary element method for solving numerically the Cauchy problem for the Laplace equation”, Engineering Analysis with Boundary Elements Vol. 20, pp123-133.
- [2] S Andrieux, T N Baranger & A Ben Abda, (2006) “Solving Cauchy problems by minimizing an energy-like functional”, Inverse Problems Vol. 22, pp115-133.
- [3] J.Hadamard, (1923) *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven.
- [4] V.Isakov, (1998) *Inverse Problems for Partial Differential Equations*, Applied Mathematical Sciences, Springer, New York.
- [5] L.E.Payne, (1975) *Improperly Posed Problems in Partial Differential Equations*, Philadelphia, SIAM.
- [6] M.M.Lavrentiev, (1967) *Some improperly posed problems of Mathematical*, Springer-Verlag Berlin Heidelberg, New York.
- [7] T.Tarantola, (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM.
- [8] A.Nachoui, (2004) “Numerical linear algebra for reconstruction inverse problems”, Journal of Computational and Applied Mathematics, Vol. 162, pp147-164.
- [9] R.V. Kohn & A.McKenney, (1990) “Numerical implementation of a variational method for electrical impedance tomography”, Inverse Problem, Vol. 6, pp389-414;
- [10] R. Lattés & J.L. Lions, (1967) “*Méthode de Quasi-réversibilité et Applications*”, Paris, Dunod.
- [11] M L. Bourgeois, (2005) “A mixed formulation of quasi-reversibility to solve the Cauchy problem for Laplace's equation”, Inverse Problems, Vol. 21, pp1087-1104.
- [12] M.V. Klibanov & F.Santosa, (1991) “A computational quasi-reversibility method for Cauchy problems for Laplace's equation”, SIAM J.Appl. Math, Vol. 51, pp1653-1675.
- [13] A. Cimetière, F. Delvare, M. Jaoua & F. Pons, (2001) “Solution of the Cauchy problem using iterated Tikhonov regularization”, Inverse Problems, Vol. 17, pp553–570.
- [14] V. A. Kozlov, V.G .Mazyra & A.V. Fomin, (1991) “An iterative method for solving the Cauchy problem for elliptic equation”, Comput. Math. Phys. Vol. 31, pp45-52.
- [15] M. Jourhmane & A. Nachoui, (2002) “Convergence of an alternating method to solve the Cauchy problem for poisson's equation”, Applicable analysis, Vol.81, pp1065-1083.
- [16] L. Marin, L. Elliott, P.J. Heggs, D.B. Ingham, D. Lesnic & X. Wen (2003) “An alternating iterative algorithm for the Cauchy problem associated to the Helmholtz equation”, Comput. Methods Appl. Mech. Engrg., Vol.192, pp709-722.
- [17] H. W. Engl & A. Leitão (2001) “A Mann iterative regularization method for elliptic Cauchy problems”, Numer. Funct. Anal. and Optimiz., Vol.22, pp861-884.
- [18] M. Jourhmane, D. Lesnic & N.S.Mera, (2004) “Relaxation procedures for an iterative algorithm for solving the Cauchy problem for the Laplace equation”, Engineering Analysis with Boundary Elements, Vol.28, pp655-665.
- [19] M. Jourhmane & A. Nachoui, (1999) “An alternating method for an inverse Cauchy problem”, Numerical Algorithms, Vol. 21, pp 247–260.
- [20] C.tajani, J;Abouchabaka, O.Abdoun, (2011) “Numerical simulation of an inverse problem: Testing the influence data”, in.proc.of the 4th International Conference on Approximation Methods and Numerical Modeling in Environment and Natural, pp29-32.
- [21] M.Essaouini, A.Nachoui & S.El Hajjia,(2004), “ Numerical method for solving a class of nonlinear elliptic inverse problems”, J.comp.appl.math, Vol.162, pp165-181.

Authors

C. Tajani, was born in Morocco in 1978. He prepared his thesis in the faculty of sciences in Kenitra, Morocco. His research interests include the Numerical Simulation of inverse problems, Shape Optimization problem. In addition to Genetic algorithm.



J. Abouchabaka, is a professor in the faculty of science, Dept. Computer science, Ibn Tofail university, Kenitra, Morocco.

A COMPREHENSIVE PERFORMANCE ANALYSIS OF ROUTING PROTOCOL FOR ADHOC NETWORK

Sachin Dahiya¹, Manoj Duhan², Vikram Singh³

¹Associate Professor, Haryana Institute of Technology, Asodha, Haryana, India

²Professor, Deptt. of ECE, DCRUST Murthal, Sonapat, Haryana, India

³Professor, Deptt. of Computer Science, CDLU Sirsa, Haryana, India

ABSTRACT

Routing is always a challenging task in a mobile adhoc network. Several protocols have been proposed for adhoc network with different functionality. All the protocols have been designed to determine the routes between the communicating nodes based upon different criteria. Due to the movement of the nodes the topology of the adhoc network is not fixed and it poses a challenge for the routing protocol to provide a path from the source to destination in such a dynamic scenario. High mobility combined with the limited energy supply of the nodes affects the performance of the routing protocol. In this paper performance of three major routing protocols DSDV, DSR and AODV have been evaluated in different network condition. Protocols are also tested for energy consumed during packet transmission, network lifetime, average throughput, packet delivery fraction, end to end delay etc.

KEYWORDS: DSDV, DSR, AODV, Adhoc Network

I. INTRODUCTION

Mobile adhoc network [6] is a dynamic network formed by large number of nodes. Adhoc network is an autonomous network and it works without the aid of any centralizing authority. Due to the mobility of the nodes, routing is quite a challenging task. The dynamic topology of the adhoc network leads to the frequent breakup of routes. Route failure affects the connectivity of the network. Moreover the nodes are dependent on the limited battery power. Power shortage in any node may results in the network partitioning.

There are two types of routing protocols for adhoc network, table driven and on demand routing protocol. Table driven protocols are also known as proactive routing protocol. In table driven protocol each node maintains the routing information of all the nodes in the network in advance. Routing information is periodically updated and exchanged among the neighbor. Periodic exchange of routing table information generates large number of overhead bits. DSDV[1] (Destination Sequenced Distance Vector) is an example of table driven protocol. Proactive routing refers to the condition that whenever a node has some data for a particular destination it can transmit immediately.

On the other hand reactive routing protocol determines the routes as and when it required by a node in the network. On demand route creation significantly reduces the control overhead. Routes are determined by sending route request packet to the immediate neighbors in response to the route request packet intermediate nodes or the destination replied by unicasting the route reply packet. AODV [2] & DSR [3] are the on demand routing protocols.

The remaining paper is organized as follows: In section 1 DSDV, DSR and AODV protocols have been discussed in brief. Section 2 describes the various metrics used for performance evaluation of routing protocols. Section 3 contains the simulation parameters for performance evaluation.

II. DESCRIPTION OF ROUTING PROTOCOLS

In this section basic features of DSDV, DSR and AODV protocols are summarized.

2.1. DSDV (Destination Sequenced Distance Vector Routing)

In this protocol every node in the network maintains the routing tables in which the routing information of all possible routes is stored. DSDV protocol is based on the Bellman- ford routing protocol. DSDV protocols provide a unique path to the desired destination. The selected path will be the shortest route among the entire possible routes. DSDV protocol routing table contains information about all the receivers. The entries present in the routing tables are marked with sequence number. Routes with the latest sequence number are preferred for making the forwarding direction. In order to reduce the amount of control overhead the two types of packets are utilized. "Full dump packets" are utilized to convey all the available Information where as the second type of packet called Incremental packet are sent to convey information that has changed after the last full dump packet. In spite of utilizing the incremental packet the DSDV protocols still generates a large number of overhead bits making it less suitable for larger network.

2.2. AODV (Adhoc on Demand Distance Vector Routing)

AODV routing protocols was designed and developed for MANET. AODV basically is an improvement over DSDV. In DSDV the route information of all the possible routes have stored in routing table of a node. AODV minimizes the number of route broadcast by creating the routes on demand basis. AODV provide loop free route and maintain a sequence number which if increased every time when ever a change is detected in the environment. Route discovery process requires the execution of the following steps:-

- a) If a node has some data to send, it will broadcast a route request message to its neighbors.
- b) Intermediate nodes received the RREQ message and store the address of the source node in the routing table so that if duplicate RREQ is received in future may be discarded.
- c) If any node has a current route to the destination lies send a response to the source in the form of RREP message. The node may be the destination node.
- d) During the reverse movement of RREP, nodes in this path will store the forward route entries for setting up of forward route to the destination. The newly determined route is associated with a timer which amounts to the time of deletion of the route whenever the route is no longer in use. After the route establishment the route maintenance phase is required whenever a movement of nodes along the active path may be detracted. Route error message (RRER) is sent to the source node. This message is generated by the node which is near to break point. After getting RRER message the source node will reinitiate the route discovery procedure.

2.3. DSR (Dynamic Source Routing)

Dynamic source routing protocols (DSR) is a reactive on demand routing protocols. In this protocol mobile node utilizes the concept to route cache to store the information of routes. If a mobile node has some data to send, it first checks its cache to determine the route availability. If an alive route is found then the node will send its data along the already existing route. If not, the node will start its basic route discovery mechanism by broadcasting route request RREQ message. After recording this route request packet the intermediate node will search the route in the cache based upon the information in route request packet. If the route to the destination is available they replied back to the source node. If the node is not destination node then it add its own address and rebroadcast the message to its neighbors. This process will ends at the destination nodes.

III. PERFORMANCE METRICS

3.1. Packet Delivery Fraction

It is defined as the ratio of the packet received by the destination successfully to the total number of packets sent by the sender.

3.2. End to End Delay

It is defined as the time taken by a packet in reaching to destination from the time it is sent by the sender. It includes all kinds of delays like queuing delay propagation delay etc.

3.2. Normalized Routing Overhead

Normalized routing overhead is the number of routing packets transmitted per data packet received at the destination.

3.3. Throughput

It is defined as total number of packets received by the destination.

3.4. Energy Consumption per Packet [4]

It is defined as the total energy balance to the total number of packets delivered to the destination.

3.5. Network Life Time [4]

Network life time may be defined as the time when the first node die due to battery failure.

IV. SIMULATION SCENARIO

DSDV, DSR & AODV protocols have been tested using the network simulator [5] (NS2). It is discrete event simulator. This simulator software is widely acceptable among the research community. Implementations of DSDV, DSR & AODV protocol are inbuilt in ns2. Protocols are tested for energy consumed during packet transmission, network lifetime, average throughput, packet delivery fraction, end to end delay etc.

4.1. First Scenario

In the first case we have calculated average throughput, packet delivery fraction, normalized routing overhead and end to end delay. In this test following parameters have utilized as listed in table 1. Pause time indicates the mobility of the nodes. Pause time corresponds to the zero seconds indicates that mobile node is continuously moving without resting at a point. Whereas 100 seconds pause time corresponds to zero mobility.

Table 1. Parameters

Parameters	Value
Number of Nodes	50
Pause Time	0,10,30,50,70,100 sec
Number of source	10
Rate	4 packets/sec
Area	700*700
Speed	20 m/sec

Figure 1 shows that the Average throughput of the DSR and AODV protocols is much higher than DSDV. Effect of mobility on the throughput of DSR & AODV is negligible. Throughput of DSDV increases with the value of pause time. DSDV performs better at low mobility levels

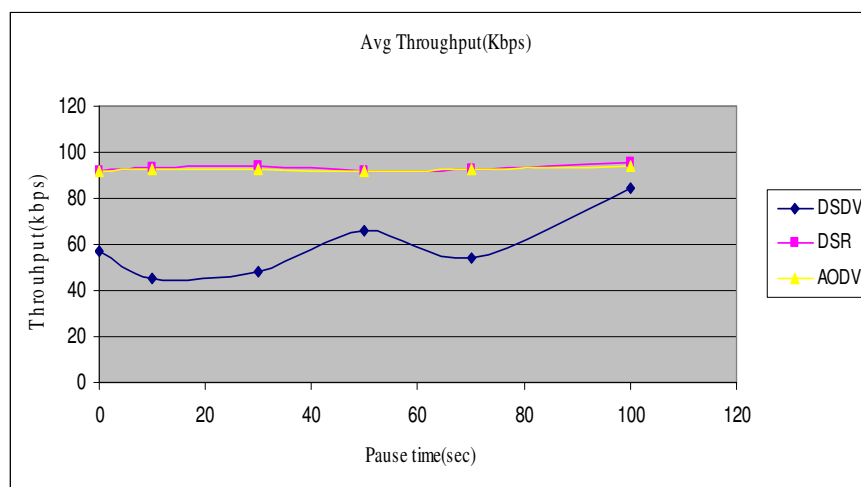


Fig 1 Throughput Vs Pause time

Figure 2 shows that the end to end delay of AODV is largest than DSR & DSDV. It is because at higher mobility, links are frequently broken. Since routes are available in the cache of DSR, hence Route discovery procedure is less required than the AODV in which routes are discovered whenever a change in the topology is detected. In DSDV routes are known well in advance, hence delay is less.

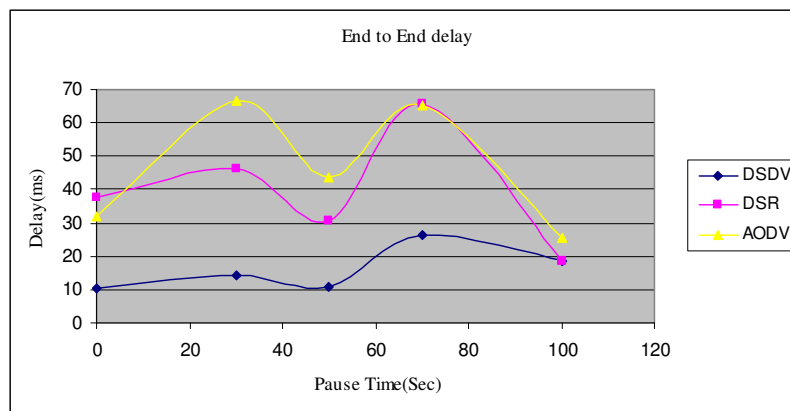


Fig 2 End to End Delay Vs Pause time

Routing overhead is also high for AODV than DSR & DSDV at high mobility level as shown in figure 3. Since frequent route discovery takes place in the AODV which leads to the generation of large number of control packets. DSDV protocol periodically exchange the routing table among the neighbor. Therefore it is also generating the significantly control overhead. Route availability in the cache makes the DSR most suitable candidate as far as this metric is concerned.

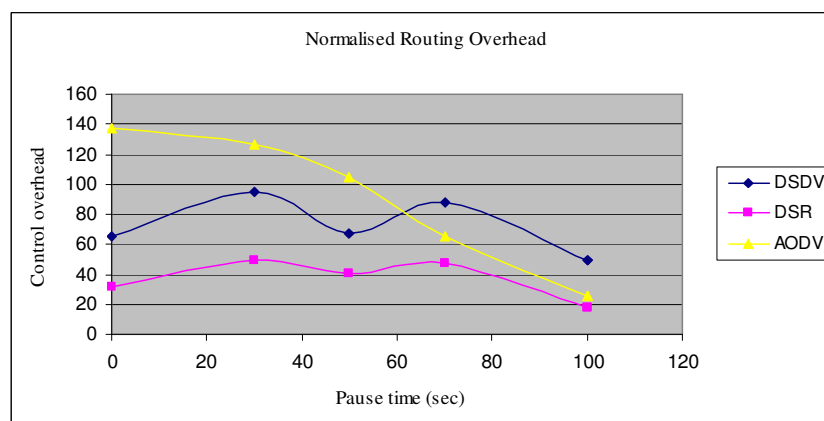


Fig 3 Normalized routing overhead Vs Pause time

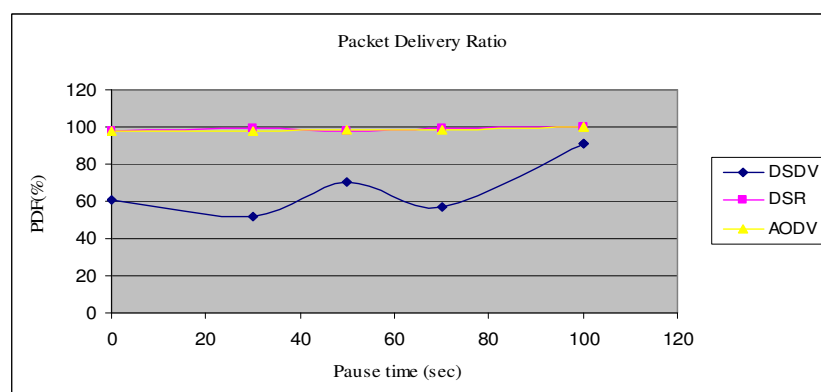


Fig 4 Packet Delivery Ratio Vs Pause Time

Packet delivery ratio of AODV and DSR are again greater than the DSDV as shown in figure 4. Above result indicates that throughput and packet delivery fraction of DSR & AODV is better than DSDV. AODV provides the high delay and generates huge control over head. Moderate delay and small overhead coupled with high throughput and packet delivery fraction shows that the DSR is the best choice among the three protocols for the adhoc network.

4.2. Analysis on the basis of Energy Metrics

Further the three protocols have been tested for energy metrics. The metrics considered are Energy consumed per packet and for network life time. The following parameters are utilized.

Table 2 Energy Parameters

Parameters	Value
Number of Nodes	50
Pause Time	0,30,50,70,100 sec
Number of source	10
Rate	4 packets/sec
Area	700*700
Speed	5 m/sec
Energy	3J

Fig 5 shows that the energy consumption of DSR is lower than AODV & DSDV. DSR consumes less energy than DSDV and AODV because of its route caching strategy. DSDV consumes more energy because of its periodic exchange of routing information. At higher level of mobility AODV and DSR have the same behavior where as at low level of mobility the performance of DSR is better than AODV. This is due to the fact that at higher level of mobility probability of choosing the stale route from the cache increases.

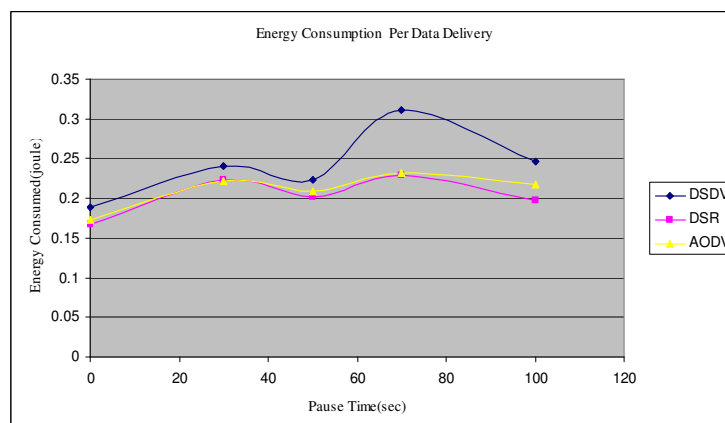


Fig 5 Energy Consumption Vs Pause Time

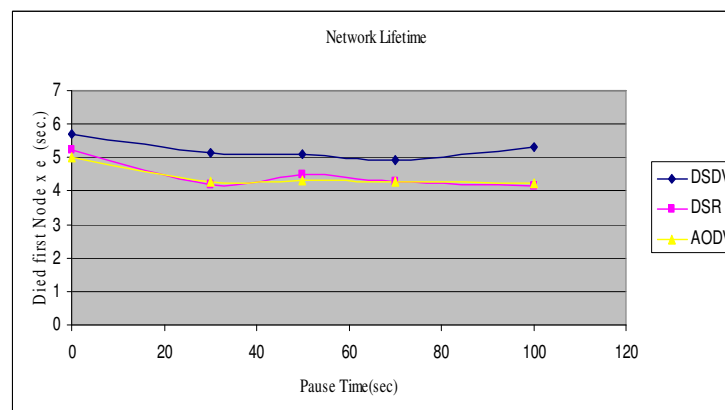


Fig 6 Network Life Time Vs Pause Time

Figure 6 shows the network lifetime of three protocols. As already stated that the network life time[4] defined in terms of the time when the first node died due to the battery failure. DSDV has the highest network life time. Since DSR depends upon the cache entries so it may be possible that some of the nodes will be over utilized than others. Moreover the mobility of the nodes has negligible impact on the lifetime. Performance of on demand routing protocols is poor as compared to the proactive protocols for this metric.

Figure 7 shows the pattern of energy consumption with node speed. There are smaller variations in the consumption pattern with speed. Energy consumption of DSDV is greater than DSR and AODV. At lower value of node speed i.e. at 1m/sec energy consumption of AODV and DSDV is almost equal. DSR performs better at low speed. It utilizes the route from the cache. At high speed the links are broken more frequently due to which caching of route will not provide a great help, rather it increases the retransmission of packets due to forwarding on the expired routes.

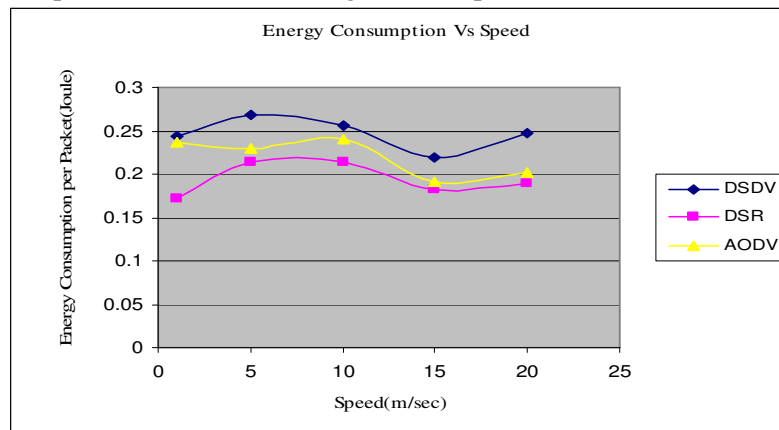


Fig 7 Energy consumption Vs node speed at Pause time 10 seconds

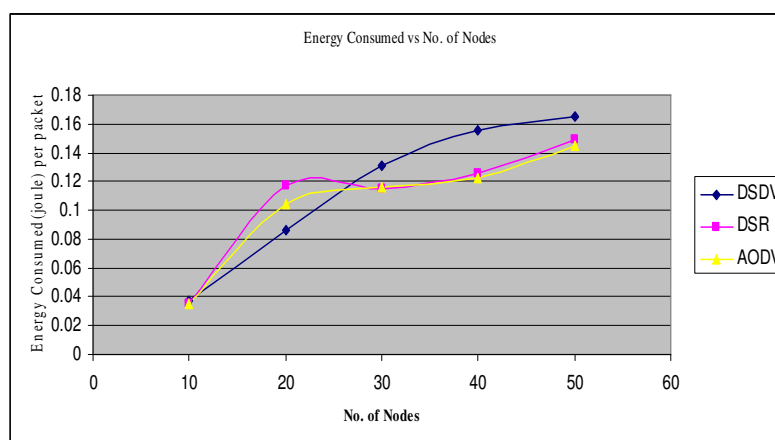


Fig 8 Energy consumption vs No. of node at Pause time 10 seconds

In sparse network all the three protocols consumes equal energy. As the network becomes denser than the energy consumption of DSDV increases because the routing information increases with the increasing number of nodes and hence circulation of routing table between nodes consumes more energy as shown in figure 8. On the other hand in AODV and DSR energy consumption also increases with increasing number of nodes. In large networks route maintenance phase consumes more energy than route creation phase. Figure 8 suggest that the DSDV is not suitable for large networks.

V. CONCLUSION

In this paper the DSDV, AODV, DSR protocols are tested in different scenario. On demand protocols AODV and DSR performs better in terms of average throughput, packet deliver fraction at high node speed. AODV has the longest end to end delay because it always establish route on demand basis. However cache mechanism of DSR reduces the end to end delay. DSDV because of its proactive

nature provides the minimum delay. AODV protocol generates largest control overhead among the three protocols. DSR consumes less energy than AODV & DSDV. DSDV provides the maximum lifetime than other protocols because DSDV utilizes all the nodes efficiently than others. Energy consumption of DSR is less than AODV & DSDV with varying node speed. Similarly DSR performs well in high density networks. The above analysis suggests that DSR is most suitable Protocol in different network condition. DSR energy consumption is low, provides good throughput, High Pdf value Moderate delay. In future more research can be done in order to increase the life time of the nodes by energy efficient load balancing techniques.

REFERENCES

- [1] Perkins C. E. and Bhagwat P., "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", SIGCOMM 94 -8/94 London England UK @ 1994 ACM 0-89791 - 682-4/94/0008.
- [2] Perkins C. E. and Royer E. M., "Ad-Hoc On-Demand Distance Vector Routing, Mobile Computing Systems and Applications," Proc. IEEE Workshop Mobile Computing Systems & Applications (WMCSA '99), pp. 90-100, 1999.
- [3] D. B. Johnson and D. A. Maltz, The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR), IETF Internet draft, July 2004, <http://www.ietf.org/internet-drafts/draft-ietf-manetds10.txt>.
- [4] Alemneh Adane, "Active Communication Energy Efficient Routing Protocol of Mobile Ad Hoc Networks (MANETS)" school of Graduate Studies, Faculty of Technology, Electrical & Computer Engineering Deptt. , Addis Ababa University, Ethiopia, Apr 2008.
- [5] NS-2 Network Simulator, <http://isi.edu/nsnam/ns/>
- [6] Magnus Frodrigh and Per Johansson & Peter Larson, Wireless ad hoc networking- The art of networking without a network.

Authors

Sachin Dahiya is presently working as an Associate Professor in Department of Electronics & Communication Engineering at Haryana Institute of Technology, Asodha, Bahadurgarh, Haryana. He has earned his Masters in Technology from NIT Kurukshetra. Presently he is doing PhD in Adhoc network from DCRUST Murthal, Haryana



Manoj Duhan is Professor & Chairman in ECE Dept at Deenbandhu chhotu ram University of Science & Technology Murthal, Sonapat, since 2008. He is also having the charge of Director Weekend Courses at DCRUS&T Murthal. He was Reader & Chairman, ECE Dept. GJUS&T Hisar from 2004 to 2008. Prior to this he was lecturer at Technological Institute of Textile & Sciences, Bhiwani. He obtained his B.E degree from COET Jalgaon, NMU in 1997. He was Gold Medalist in B.tech. He received his M.Tech degree from dept. Of Electronics, NIT Kurukshetra in 2001, followed by Ph.D degree in 2005 from Technological Institute of Textile & Sciences, Bhiwani, and M.D.U Rohtak. His areas of interest are Reliability, DSP, wireless Communication, Microprocessor. He is life member of IAENG, ISTE, ISITA, MCU and SSI. He was awarded "Best Lecturer Award" by Governor of Haryana in 1998 at SBMNEC Rohtak. In July 2007 he was awarded "Shiksha Ratan Award" by India International Friendship society, New Delhi. He has published more than seventy research papers in National & International Conferences & Journals.



Vikram Singh is presently working as a Professor of Computer Science and Dean, Faculty of Physical Sciences at Chaudhary Devi Lal University, Sirsa, Haryana (India). He has earned his Masters in Physics, Masters in Computer Science and Doctorate in Computer Science – all – from Kurukshetra University, Kurukshetra. Earlier, he has worked as a Lecturer with Department of Computer Science & Applications, Kurukshetra University, Kurukshetra. He has also worked as a System Programmer on a World Bank Project for four and a half years. Areas of his research interest include simulation and modeling, electronic governance, online social media technologies and data mining.



COMPARISON OF GA AND LQR TUNING OF STATIC VAR COMPENSATOR FOR DAMPING OSCILLATIONS

Nuraddeen Magaji, Mukhtar F. Hamza, Ado Dan-Isa
Department of Electrical Engineering, Bayero University, PMB 3011 Kano, Nigeria.

ABSTRACT

Static VAR Compensator (SVC) is added to the excitation system of a generator in order to enhance damping during low frequency oscillations. In this article Linear Quadratic Regulator (LQR) tuning approach and Genetic Algorithm (GA) are used to obtain supplementary controller parameters for oscillation damping. The performances of SVC without controller, SVC with LQR controller and SVC with GA controller are compared. Eleven bus two area four machine system is used for the study. The performance of the GA controller is found to be better than the LQR controller and without controller under different operating conditions. A comparison between the effect of SVC alone and the proposed GA and LQR is made.

KEYWORDS: *Static VAR Compensator (SVC), GA, LQR, Power system oscillation*

I. INTRODUCTION

Power system oscillations are as a result of lack of sufficient damping torque at the generators rotors. This situation may happen as a result of heavy load in the lines, weak interconnection, high gain excitation systems etc[1]. The oscillation of the generators rotors cause the oscillation of other power system variables such as transmission line active and reactive powers, bus voltage, bus frequency, etc. Depending on the number of generators involved, the frequency of the oscillation is usually between 0.1 and 2Hz [2]. There are several types of oscillations: Local mode, Inter area mode, Control mode and Torsion mode[3,12]. Many devices can be used as damping controllers for power system oscillation such as flexible ac transmission system (FACTS) devices, power system stabilizers (PSS), high voltage dc (HVDC) links, static var compensators, etc. PSS is applied on selected generators to damp the local mode oscillation but for inter-area mode oscillation a supplementary controller is applied to the SVC devices. In most cases the design of these controllers are based on a linearised model which, for wide range of operating points and under large disturbances, will not provide satisfactory performance[1]. The main objective of this paper is to compare the performances of supplementary SVC controller, genetic-algorithm-based controller and LQR on oscillation damping. SIMULINK is used for the simulations in MATLAB environment.

The remaining sections of this paper are arranged as follows. Section 2 discusses static var compensating model. The relevant equations modeling the SVC current and reactance are derived. In Section 3 a small overview of Genetic Algorithm (GA) is given. Similarly, overview on LQR is given in Section 4. Details of the actual power system model and the results are presented in Section 5. The paper is concluded in Section 6.

II. STATIC VAR COMPENSATOR MODEL

SVC consists of reactors and capacitors, with thyristor control valves which are connected in parallel with a fixed capacitor bank. It is connected to the transmission line via a transformer as shown in fig.

1. The equivalent circuit of the SVC is shown in fig 2. Its position in the transmission line depends on the primary usage (i.e. midpoint of long transmission line or near load center) [2]. The voltage of a bus is regulated by controlling or adjusting the equivalent reactance of the SVC[2]. This constitutes the main function of the SVC. Being an adjustable reactance device, the SVC is capable of performing both inductive and capacitive compensation.

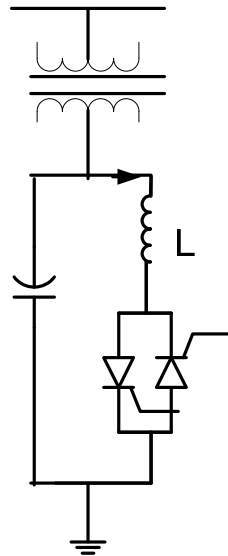


Fig. 1 SVC MODEL

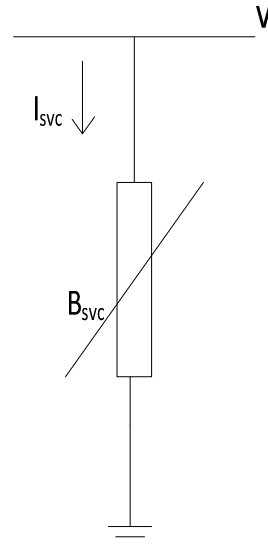


Fig 2 SVC equivalent circuit

2.1 Reactive Power Flow Model

The thyristors conduct during the alternating half-cycles of the voltage waveform depending on the firing angle α , which is measured from a zero crossing of the voltage. Full conduction is obtained with a firing angle of 90° . The current is essentially reactive and sinusoidal. Partial conduction is obtained with firing angles between 90° and 180° . Firing angles less than 90° are not allowed as they produced asymmetrical current with a dc component. The effect of increasing the firing angle is to reduce the fundamental harmonic components of the current. This is equivalent to an increase in the inductance of the reactor, reducing its reactive power as well as its current[6]. The conduction angle σ can be defined as a function of the firing angle α .

$$\sigma = 2(\pi - \alpha) \quad (1)$$

The instantaneous current in the Thyristor-Controlled Reactor (TRC) is giving by

$$I(t) = \begin{cases} \frac{\sqrt{2}V}{X_L} (\cos \alpha - \cos \omega t) & \alpha < \omega t < (\alpha + \sigma) \\ 0 & (\alpha + \sigma) < \omega t < (\alpha + \pi) \end{cases} \quad (2)$$

Where ω is the supply voltage frequency, V is the voltage r.m.s applied to the TRC and $X_L = \omega L$ is a fundamental-frequency reactance of the reactor. The fundamental component is found by Fourier analysis and is given by

$$I_L = \frac{V}{X_L} \frac{\sigma - \sin \sigma}{\pi} [A] r.m.s \quad (3)$$

We can write equation (3) as

$$I_L = B_L(\sigma) V \quad (4)$$

Where $B_L(\sigma)$ is an adjustable fundamental-frequency susceptance controlled by the conduction angle according to the following law

$$B_L(\sigma) = \frac{\sigma - \sin \sigma}{\pi X_L} \quad (5)$$

This can also be written as a function of the firing angle as

$$B_L(\alpha) = \frac{2(\pi - \alpha) + \sin 2\alpha}{\pi X_L} \quad (6)$$

The maximum value of the variable susceptance is $1/X_L$, obtain with $\sigma = 180^\circ$ ($\alpha = 90^\circ$) and the current on the reactor is maximum. The minimum value is zero, obtained with $\sigma = 0^\circ$ ($\alpha = 180^\circ$). This control principle is called phase control [7]. The equation of the equivalent susceptance is given by

$$B(\sigma) = B_C - B_L(\sigma) \quad (7)$$

Where B_C is the capacitor susceptance, B_L is the reactor susceptance. The reactive power injected at the SVC node is then given by

$$Q(\sigma) = -(B_C - B_L(\sigma))V^2 \quad (8)$$

2.2 System Dynamic Model

The SVC model with supplementary controller is shown in fig.3. In this model, a total reactance B_{SVC} is assumed and the following differential equation holds [2].

$$\dot{B}_{SVC} = (K_R(V_{ref} - V_T) - B_{SVC})/T_R \quad (9)$$

The regulator has an anti-windup limiter, thus the reactance B_{SVC} is locked if one of its limits is reached and first derivative set to zero[2]. The supplementary input V_{ref} is to maintain acceptable voltage at the SVC bus. TCR of 150MVar is connected in parallel with fixed capacitor of 200MVar correspond to a limit of 2.0pu to -1.5pu at 1.0pu voltage.

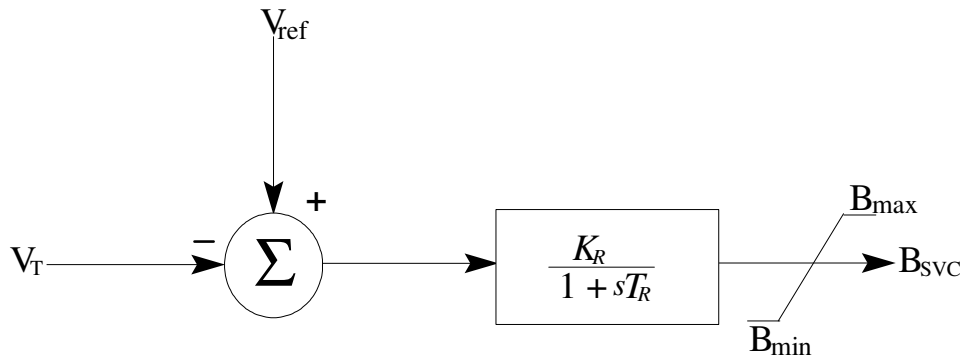


Fig. 3: SVC model with supplementary controller

III. GENETIC ALGORITHM

Genetic Algorithm (GA) is a heuristic search technique based on the evolutionary ideas of natural selection. It is usually used to solve optimization problem by random search [3]. GA is not restricted by difficult mathematical model and is flexible enough for almost all types of design criteria. GA can easily be integrated with already developed analysis and simulation tools [1].

There are several methods to select parents from the old population, and there are different GA methods that can be used for different selection method. Encoding process is necessary when GA is used to solve an optimization problem (i.e. represent the solution in a string form). In most cases the binary encoding method is used. Crossover and mutation which are the standard genetic operators are operating on the string to search for optimal solutions. The solution to the SVC tuning problem can easily be encoded as binary string.

The most important operator in GA is crossover, and it is applied with probability between 0.6 to 0.9. It takes two strings from the old population and exchange some contiguous segment of their structure to form two offspring. There are many types of crossovers in GA. Another important operator in GA is mutation. In a binary encoded string the mutation operator randomly switch one or more bits with some small probability which is typically between 0.001 to 0.09 [9].

The genetic algorithm toolbox uses MATLAB matrix function to build a set of versatile tools for implementing a wide range of GA methods. The GA toolbox is a collection of routines, written mostly in script files (m-files), which implement the most important function in GA[12]. Fig 4 shows the simple GA flow chart.

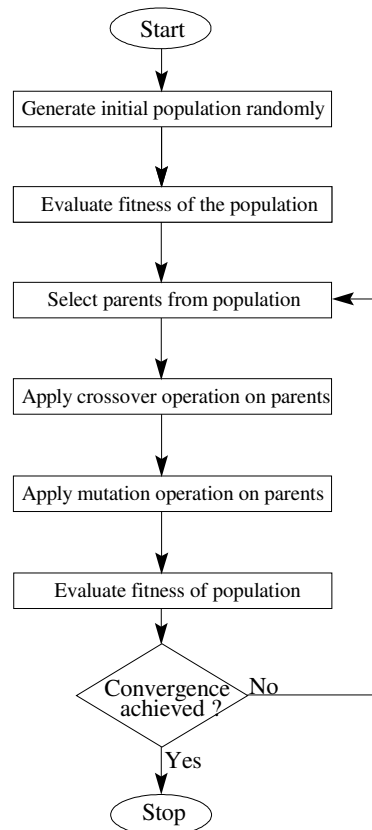


Fig 4 simple GA flow chart.

IV. LINEAR QUADRATIC REGULATOR (LQR)

Linear quadratic regulator (LQR) provides an optimal control law for a linear system. It's a control strategy based on minimizing a quadratic performance index [3, 4]. LQR design problem is that of finding a control input u that minimizes the quadratic cost function J , where

$$J = \int_0^{\infty} (X^T Q X + U^T R U) dt \quad (10)$$

Q and R are the state and control weighting matrices and are always square and symmetric. It is required that Q be positive semi-definite and R be positive definite. Given a system in state space representation

$$\dot{x} = Ax + Bu \quad (11)$$

where (A, B) is stable, the optimal control u is defined as

$$u = -Kx \quad (12)$$

The matrix K is given by

$$K = R^{-1} B^T P \quad (13)$$

The symmetric definite matrix P is the solution of the algebraic Riccati equation given by

$$PA + A^T P + Q - PBR^{-1}B^T P = 0 \quad (14)$$

The closed-loop system which has the optimal eigenvalues is given by

$$\dot{x} = A_c x = (A - BK)x \quad (15)$$

4.1 Selection of Q and R (Weight Matrices)

Weight matrices Q and R are selected such that the performance of the closed-loop system can satisfy the desired requirements. The selection of Q and R is partially related to the performance specifications, and certain amounts of trial and error are required with an interactive computer

simulation before a satisfactory design results. Now given a linear model one can use either pole placement or LQR technique to design full state feedback control given by equation (12).

The MATLAB function *lqr* can be used to find the optimal control gains for a continuous controller. One can use the *lqr* function for choosing R and Q , which will balance the relative importance of the input and state in the cost function that one is trying to optimize. The simplest case is to assume $R = I$ (identity matrix) and $Q = C^T * C$. LQR method allows for the control of all outputs. The controller can be tuned by changing the nonzero elements in the Q matrix to get a desirable response.

Finding the matrices P and Q is difficult for high order systems. In those cases, approximate solutions can be obtained by using a reduced-order model [10]. For convenience a balanced truncation and residualization method is used on the system to reduce it from 41- state to 4- state before applying LQR on it.

V. RESULTS

In this study, a two area interconnected four machine power system, shown in fig. 5, is considered. The system consists of machines arranged in two areas inter-connected by a weak tie line[10]. The location of the SVC is indicated in the diagram. The system is operating with area one exporting 400MW to area two. Network and generators data can be found in reference [5]. The SVC is treated as a variable capacitance. Simulations were performed in the Matlab using ODE45 solver with a fixed step-size of 1ms.

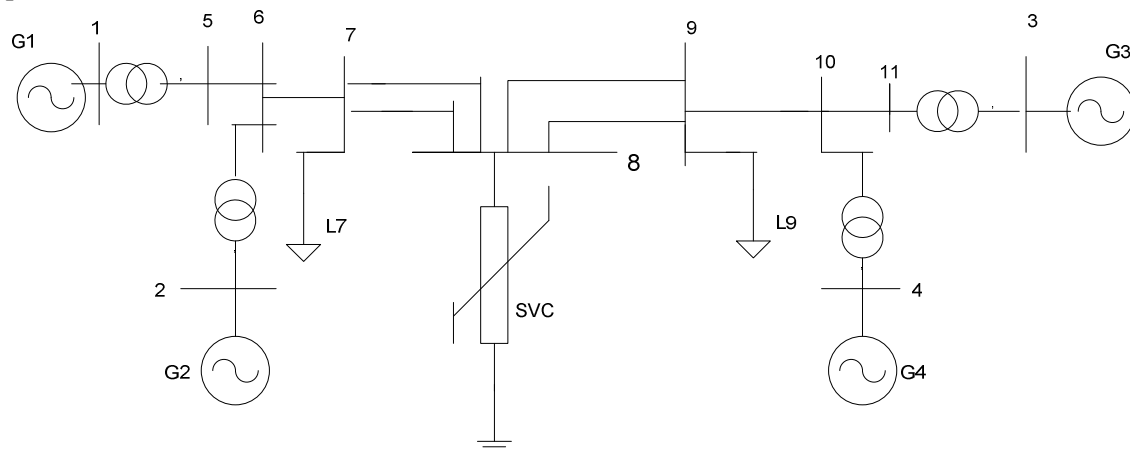


Fig.5: Two area test system with SVC

Case 1: for this case study, SVC is connected at bus 8 with normal load demand from area 2 of 400MW with all the tie line in place. Fig. 6 shows the response for the reactive power at line 7-8 without controller, with GA and with LQR controller. Fig 7 shows the speed deviation response between generator1 and generator 2 without controller, With GA controller and with LQR controller turning the SVC this figure shows the important of LQR and GA turning SVC over SVC alone. Table 1 shows the values of matrix K (LQR constants) for different cases and for different lines.

Table 1: LQR Constants.

Lines	Case 1				Case 2			
K (Line7-8)	0.00	64.09	-23.90	0.00	0.00	64.09	-23.90	0.00
K (Line8-9)	0.00	17.90	-3.68	0.00	0.00	26.90	-3.85	0.00

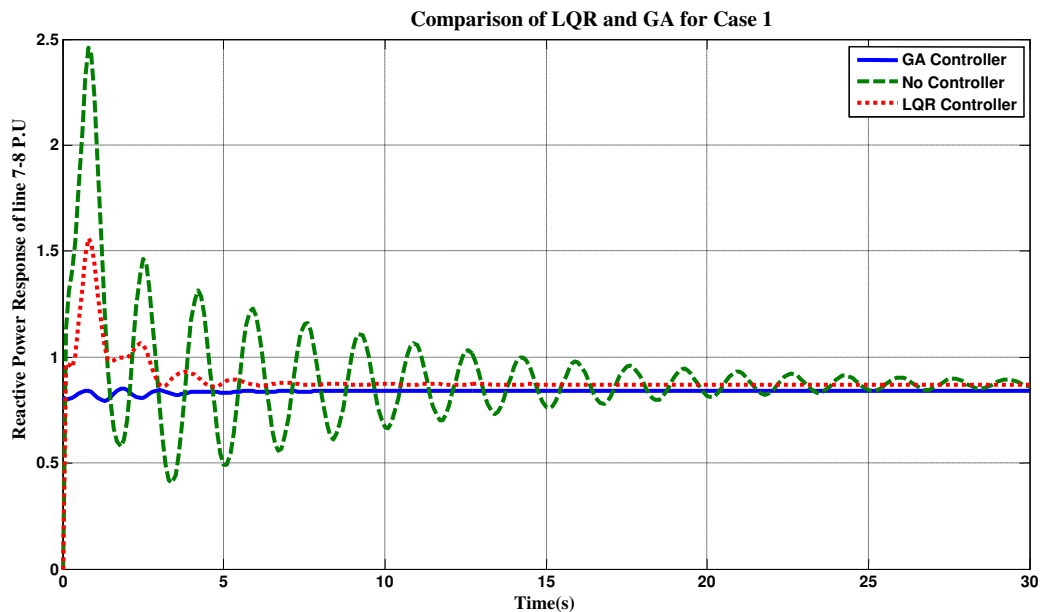


Fig 6: Reactive power response for line 7-8, case 1

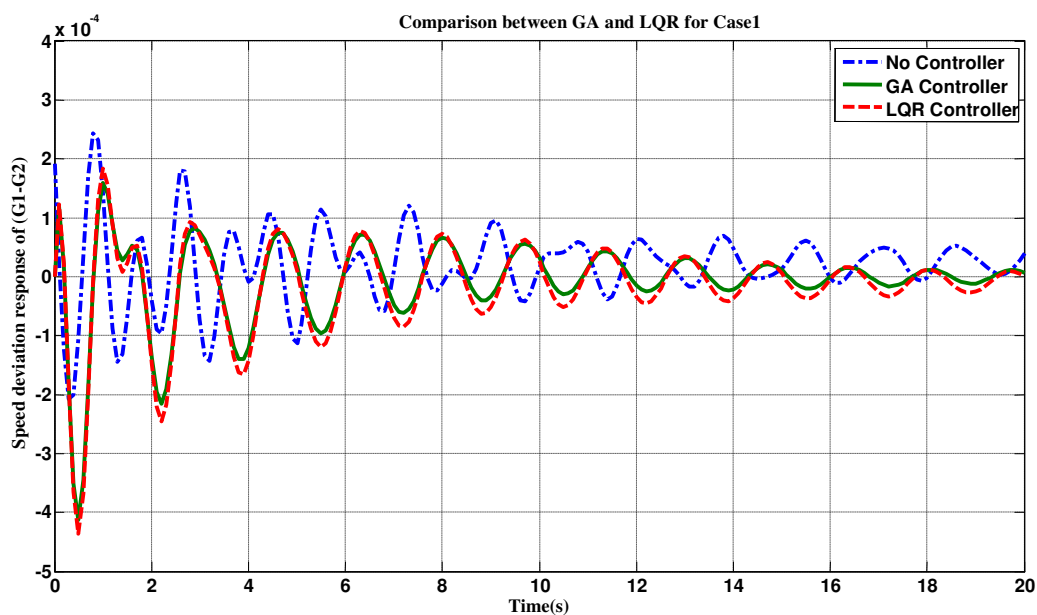


Fig. 6: Speed deviation between generator 1 and generator 2

Case 2: For this case study, a three phase faults is applied at bus 8 for a one second cleared after 1.05s with normal load demand from area 2 of 400MW and all the tie lines in place and SVC connected at bus 8. Fig. 8 shows the response for the reactive power in line 7-8 with GA Controller, without controller and with LQR controller. Fig. 9 shows the response for speed deviation between Generator 1 and 2 with GA controller, LQR controller and without controller. These figures indicate the superiority of GA controller followed by LQR controller over the SVC alone.

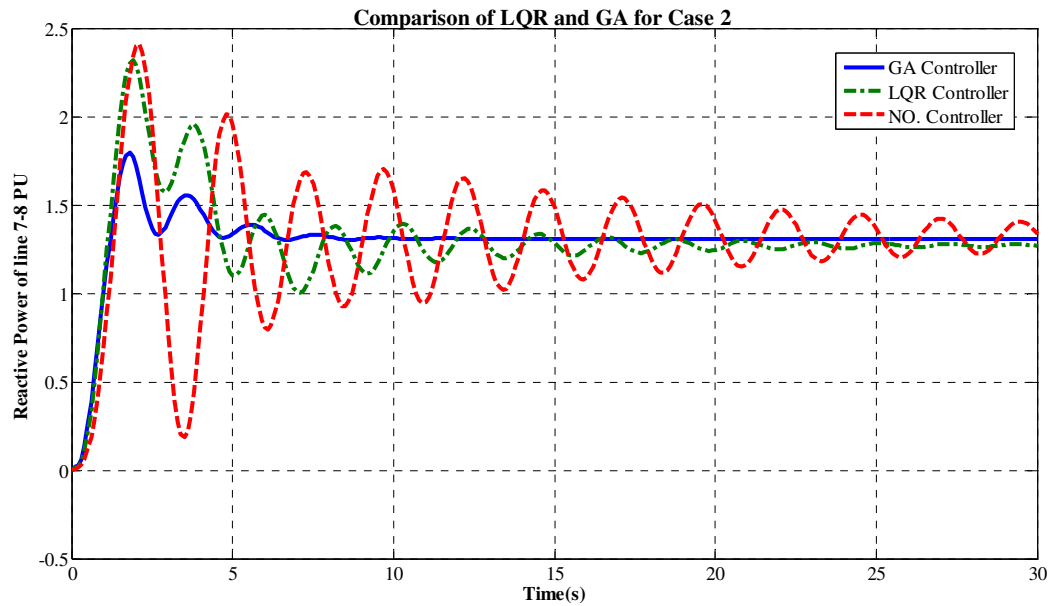


Fig 8: Reactive power response for line 7-8, case 2

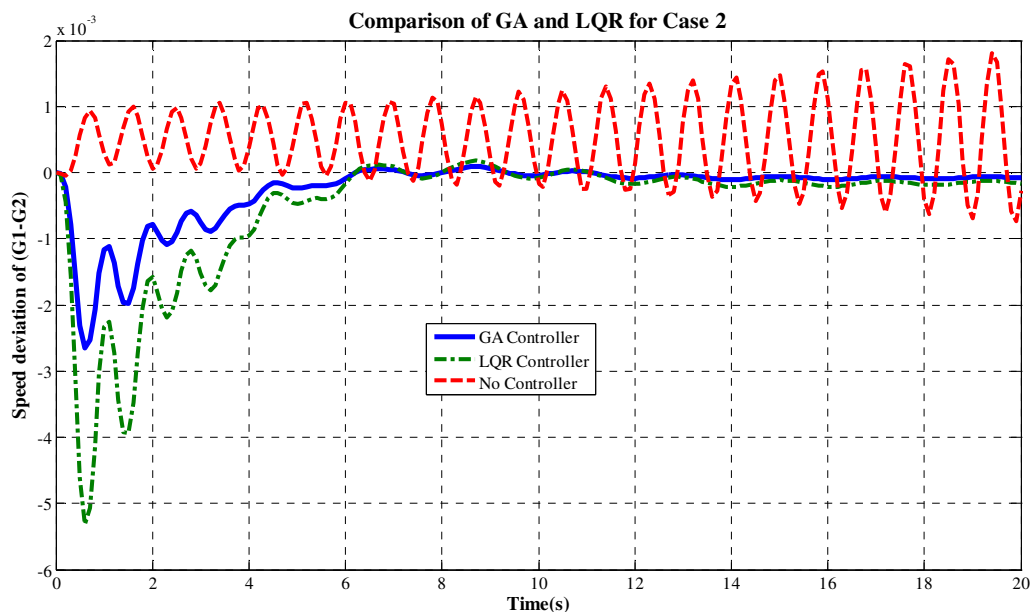


Fig 9: Speed deviation response between G1 and G2, case 2

VI. CONCLUSION

LQR tuning is proposed to obtain a supplementary controller parameter for oscillation damping. The procedure developed identify the most effective use of LQR controller parameters by considering $R = I$ and selecting symmetric matrix Q as product of transpose of output matrix and the output matrix. The simulation results show that the propose LQR under different operating conditions worked effectively and robustly for different signals.

REFERENCES

- [1] Magaji, N., Mustafa, M.W. and Muda, Z., "Power system damping using GA-based fuzzy controlled SVC device", *IEEE Region 10 Conference TENCN-2009 Singapore*, January 2009. pp1-7.

- [2] Hasanovic, A. and Feliachi, A., "Genetic algorithm based inter-area oscillation damping controller design using MATLAB", *Power Engineering Society Meeting*, **vol.3**, July 2002. pp1136-1141.
- [3] Farsangi, M.M., Nezambadi-Pour, H., Song, Y-H. and Lee, K.Y., "Placement of SVC and selection of stabilizing signal in power system, *IEEE transaction on Power Systems*, **vol 22**, No.3, 2007. pp1061-1071.
- [4] Mustafa, M.W. and Magaji, N., "Design of power oscillation damping controller for SVC devices", *IEEE 2nd International Power and Energy Conference (PECon 2008)*, Johor Bahru, December 2008. pp1329-1332.
- [5] Magaji, N. and Mustafa, M.W., "Optimal location and signal selection of SVC device for damping oscillation", *Int. Review on modeling and simulations*, **vol.2**, 2009. pp144-151.
- [6] Vachirasricirikul, S.; Ngamroo, I.; Kaitwanidvilai, "Design of robust SVC for voltage control in an isolated wind-diesel hybrid power system", *IEEE 5th Int. Conf. on ECTI-CON*, 2008. pp1053-1056.
- [7] Mustafa, M.W. and Magaji, N., "Optimal location of static VAR compensator device for damping oscillation", *American J. of Engineering and Applied Sciences*, **2**(2): 353-359, 2009.
- [8] Pal, B. and Chaudheri, B., *Robust control in power system*, Springer, 2005.
- [9] Burns, R.S., *Advanced control engineering*, Butterworth-Heinemann, 2001.
- [10] Buso, S. and Mattavelli, P., *Digital control in power electronics*, Morgan and Claypool, 2006.
- [11] Kundur P. Power system stability and control. New York, USA: McGraw- Hill; 1994.
- [12] Magaji N and Mustafa M .W. Optimal Location and Signal Selection of UPFC Device Damping Oscillations, *Int. J. Elect. Power & Energy Syst.*, 33(4) , 2011: 1031-1042.

Authors

Nuraddeen Magaji received his B Eng and M Eng. degrees from Bayero University, Kano-Nigeria in 1999 and 2005, respectively. And PhD at the Universiti Teknologi Malaysia. He is currently a lecturer at Department of Electrical Engineering, Bayero University, Kano-Nigeria. Dr Nuraddeen is also a member of IEEE. His current research interests are FACTS modeling and control, Fuzzy logic , Neural Network and power system deregulation



Mukhtar Fatihu Hamza receive his B Eng. degree from Bayero University Kano- Nigeria in 2008. presently he is a M Eng. candidate at department of electrical engineering, faculty of engineering Bayero University Kano-Nigeria. His current research interest are FACTS devices modeling and control, LQR, GA and Power system stability.



Ado Dan-Isa is a lecturer in the Department of Electrical Engineering of Bayero University, Kano, Nigeria. He received his first B.Eng. degree from the same department in 1986 and Doctor of Philosophy degree from the University of Sussex in the United Kingdom in 1996. He is a registered Engineer (COREN) and also a member of the Nigerian Society of Engineers. His current research interests are in control, soft computing, metaheuristics and computer applications.



INVERTED SINE PULSE WIDTH MODULATED THREE-PHASE CASCADED MULTILEVEL INVERTER

R.Seyezhai

Associate Professor, Department of EEE, SSN College of Engineering, Chennai, India

ABSTRACT

MultiLevel Inverter (MLI) has been recognized as an attractive topology for high voltage DC-AC conversion. This paper addresses an Inverted Sine PWM (ISPWM) technique to control the seven-level asymmetric cascaded multilevel inverter. The proposed switching technique enhances the fundamental component of the output voltage and generates lower total harmonic distortion in comparison with the conventional triangular based multicarrier PWM techniques. A detailed study of the proposed modulation technique is carried out through MATLAB /SIMULINK. Gating signals are generated using FPGA Spartan processor. A prototype of three-phase asymmetric cascaded MLI is developed to verify the theoretical results.

KEYWORDS: Asymmetric Multilevel Inverter, THD, Inverted sine PWM.

I. INTRODUCTION

Multilevel inverter is an effective and practical solution for increasing power demand and reducing harmonics of AC waveforms. Function of a multilevel inverter is to synthesize a desired voltage wave shape from several levels of DC voltages. As the number of levels increase, the synthesized staircase output waveform has more steps, approaching the desired sinusoidal waveform. They are of special interest in the distributed energy sources area because several batteries, fuel cell and solar cell can be connected through multilevel inverter to feed a load [1]. This paper presents a seven-level asymmetric cascaded multilevel inverter suited for renewable energy applications. The advantages of this topology are better output quality, lower switching dissipation, lesser number of sources and higher efficiency. The performance of multilevel inverter (such as switching loss and harmonics) is mainly decided by its modulation strategies. The modulation control schemes for the multilevel inverter can be divided into two categories fundamental switching frequency and high switching frequency PWM. Further, the high frequency PWM is classified as multilevel carrier-based PWM and multilevel space vector PWM [2]. The most popular and simple high frequency switching scheme for MLI is the Multi-Carrier PWM (MCPWM). It can be categorized into two groups: Carrier Disposition (CD) methods and Phase Shifted (PS) methods [3]. Among the carrier disposition methods, Phase Disposition (PD) PWM technique is normally employed for MLI as the carriers need minimal amounts of alteration since they are all in phase with each other [4]. This paper focuses on PD based inverted sine PWM technique for a seven-level inverter. ISPWM has a better spectral quality and a higher fundamental voltage compared to the triangular based PWM. The performance of asymmetric MLI with the proposed modulation strategy is studied through simulation studies for linear loads. A comparative evaluation between hybrid modulation and the conventional PDPWM method is also presented in terms of output voltage quality, power circuitry complexity, Total Harmonic Distortion (THD). A seven-level asymmetric MLI prototype has been developed to verify the simulation results. Section-2 discusses about the three-phase asymmetric cascaded multilevel inverter with reduced number of switches. Section -3 describes the proposed inverted sine pulse width modulation technique

highlighting its advantages compared to the conventional triangular based carrier. Section -4 discusses the simulation results for the proposed PWM technique and the results are compared with the conventional PWM. Section-5 provides the details about the prototype of the asymmetric multilevel inverter and the validation of simulation results.

II. ASYMMETRIC CASCADED MULTILEVEL INVERTER

The cascaded multilevel inverter with separate DC sources can fit many of the needs of all electric vehicles because it can use on board batteries or fuel cells to generate a nearly sinusoidal voltage waveform to drive the main vehicle traction motor. Normally, each phase of a three-phase cascaded multilevel inverter requires “n” DC sources for $2n+1$ level [5]. For many applications, multiple DC sources are required demanding long cables and this could lead to voltage unbalance among the DC sources [6]. With an aim to reduce the number of DC sources required for the cascaded multilevel inverter, this paper focuses on an asymmetric topology which uses only two DC sources to generate seven-level output. The proposed topology consists of two H-bridges as shown in Fig.1. By appropriately opening and closing the switches of H_1 , the output voltage V_1 can be made equal to $-V_{dc}$, 0 , $+V_{dc}$. Similarly the output voltage V_2 of the second bridge H_2 can be made equal to $-0.5V_{dc}$, 0 , $0.5V_{dc}$. Therefore, the output voltage of the MLI have the values of $-1.5V_{dc}$, $-V_{dc}$, $-0.5V_{dc}$, 0 , $0.5V_{dc}$, V_{dc} and $1.5V_{dc}$ as shown in Fig.2.

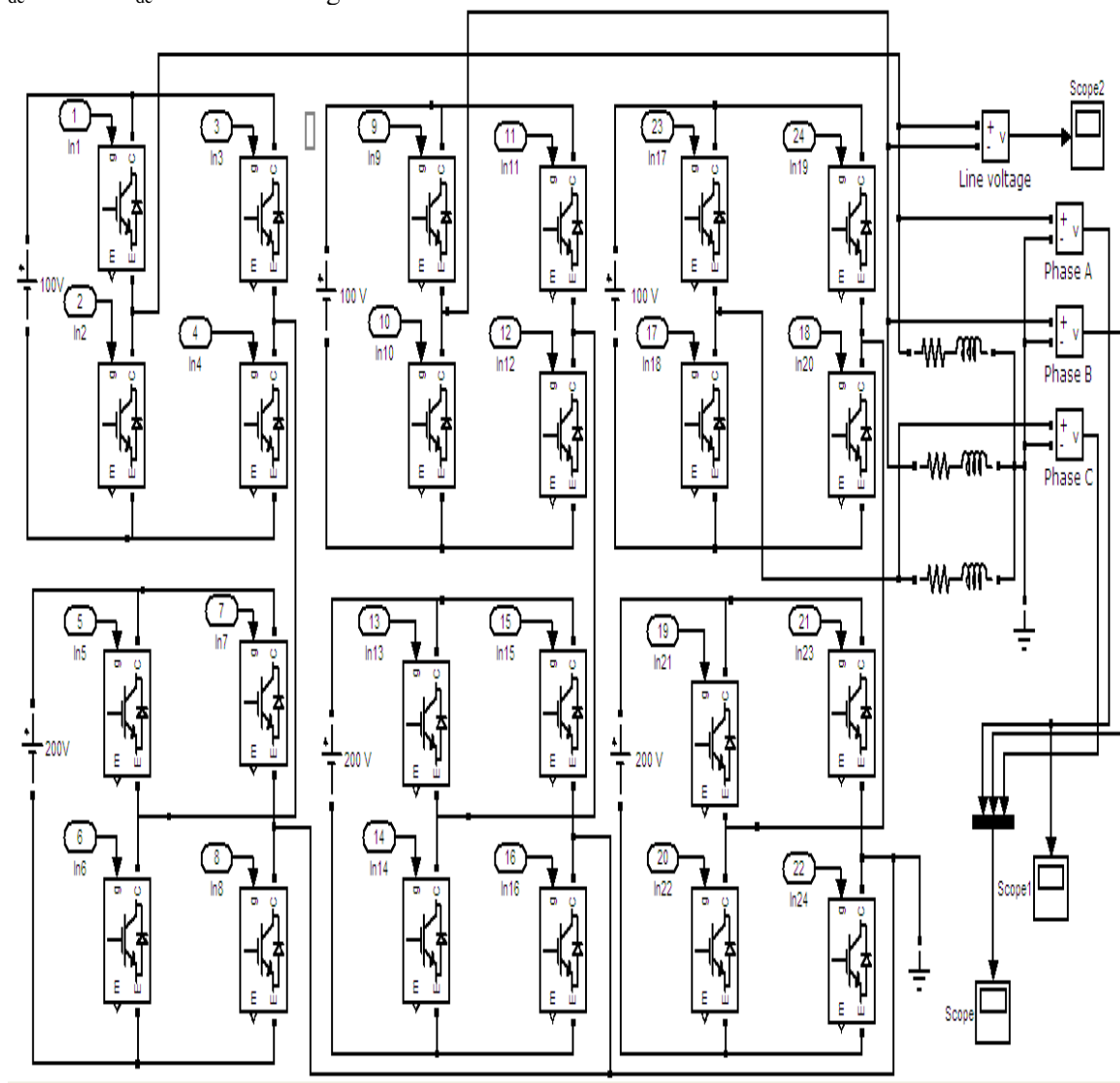


Fig.1 Three-phase Asymmetric cascaded multilevel inverter

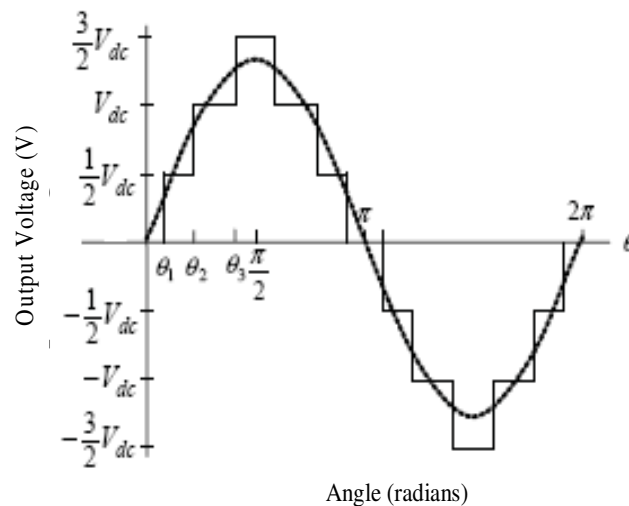


Fig .2 Seven –level equal step phase voltage waveform

III. INVERTED SINE PWM TECHNIQUE FOR MLI

The inverted sine carrier PWM (ISPWM) method uses the conventional sinusoidal reference signal and an inverted sine carrier. The control strategy uses the same reference (synchronized sinusoidal signal) as the conventional SPWM while the carrier triangle is a modified one. The control scheme uses an inverted (high frequency) sine carrier that helps to maximize the output voltage for a given modulation index [7-10]. In the gating pulse generation of the proposed ISCPWM scheme, the triangular carrier waveform of SPWM is replaced by an inverted sine waveform. In Fig.3 shows the pulse generation circuit for a single phase of the multilevel inverter in which a sine wave (modulating signal) of fundamental frequency is compared with high frequency phase disposed inverted sine carrier waves. For an ‘m’ level inverter, (m-1) carrier waves are required. The pulses are generated when the amplitude of the modulating signal is greater than that of the carrier signal.

The proposed control strategy has a better spectral quality and a higher fundamental output voltage without any pulse dropping.

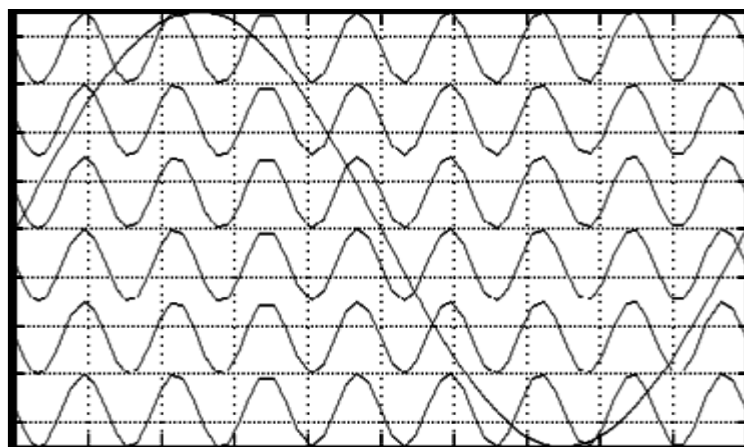


Fig.3. Reference and carrier waveforms for ISPWM technique

The advantages of ISPWM method are:

- It has a better spectral quality and a higher fundamental component compared to the conventional sinusoidal PWM (SPWM) without any pulse dropping.
- The ISCPWM strategy enhances the fundamental output voltage particularly at lower modulation index ranges

- There is a reduction in the total harmonic distortion (THD) and switching losses.
- To increase the fundamental amplitude in the sinusoidal pulse-width modulation the only way is by increasing the modulation index beyond 1 which is called over modulation. Over modulation causes the output voltage to contain many lower order harmonics and also makes the fundamental component vs. modulation index relation non-linear. Inverted sine pulse - width modulation technique replaces over modulation [11].
- The appreciable improvement in the total harmonic distortion in the lower range of modulation index attracts drive applications where low speed operation is required.

ISPWM technique causes marginal increase in the lower order harmonics, but except third harmonic all other harmonics are in acceptable level. But for three phase applications the heightened third harmonics need not be bothered [12-15]. The pulses generated using inverted sine pulse-width modulation technique is shown in Fig.4 which is used to trigger the IGBTs in a sequential manner such that the desired output is obtained.

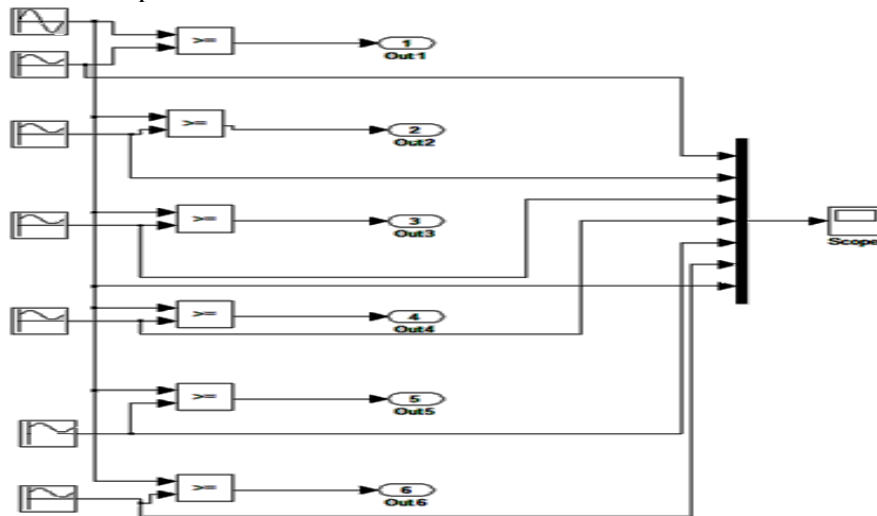


Fig. 4 Pulse generation circuit for ISPWM

The gating signals for a single phase asymmetric MLI using ISPWM is shown in Fig.5.

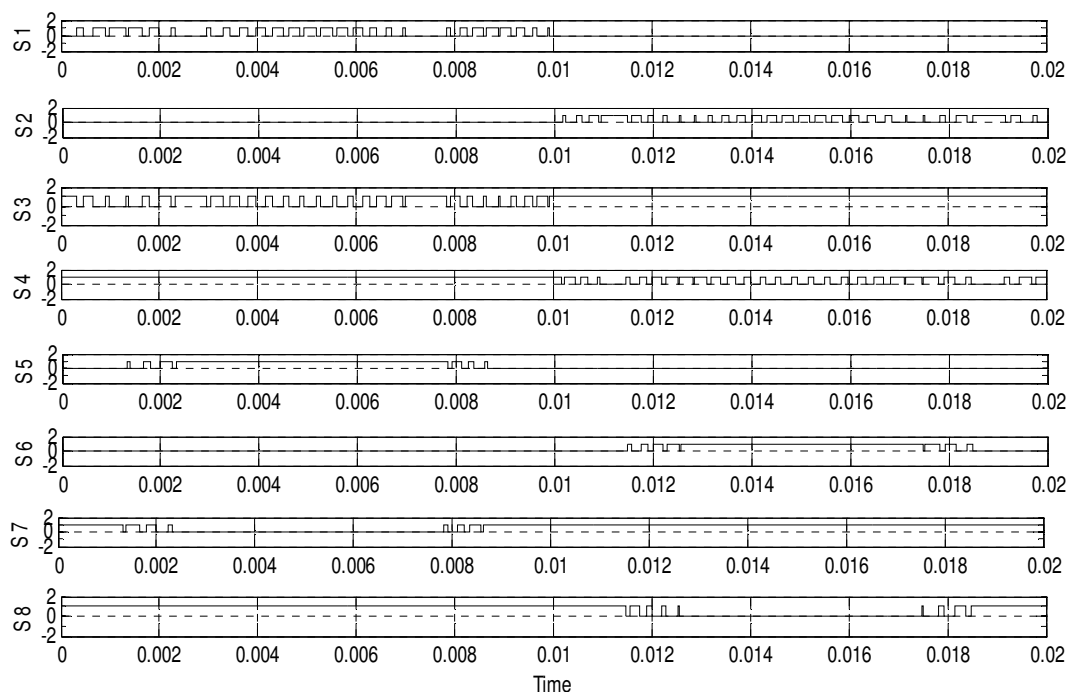


Fig.5 Pulse pattern for asymmetric MLI employing ISPWM

IV. SIMULATION RESULTS

The performance evaluation [16-17] of an inverted sine pulse-width modulated for three-phase asymmetric cascaded multilevel inverter is done using MATLAB and the phase voltage for the three phases is shown in Fig.6

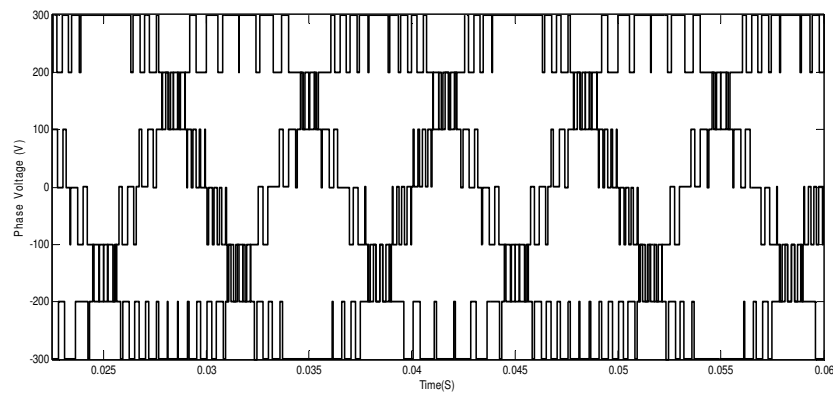


Fig.6 Phase Voltage waveform for asymmetric MLI using ISPWM

The line- line voltage waveform is shown in Fig.7

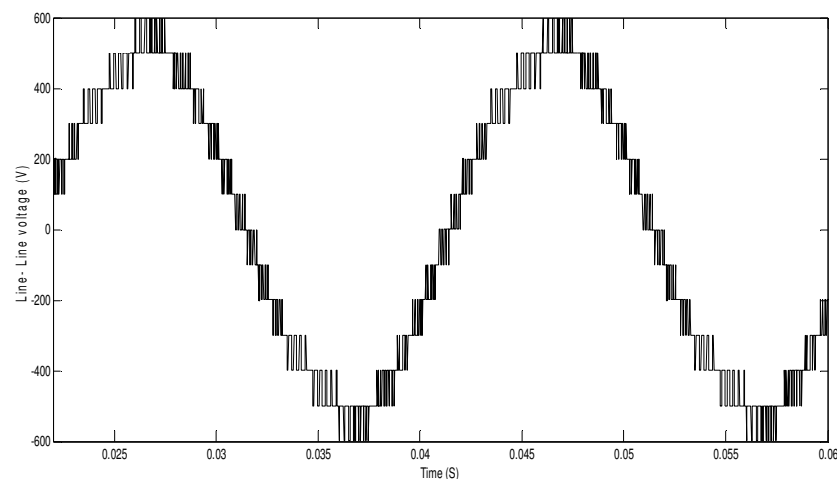


Fig.7 Line-line voltage waveform of Asymmetric cascaded MLI (ISPWM)

The line- line current waveform for asymmetric cascaded MLI is shown in Fig.8

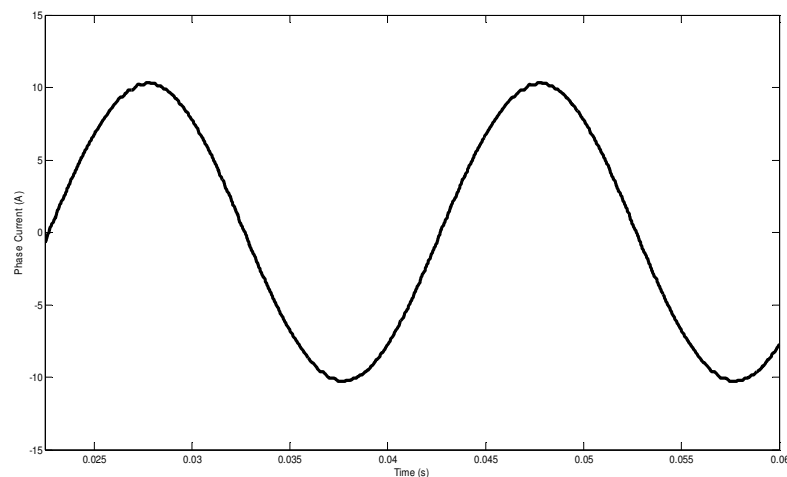


Fig.8 Line –line current waveform of Asymmetric cascaded MLI (ISPWM)

FFT spectrum for the load voltage waveform of MLI using ISPWM technique is shown in Fig.9

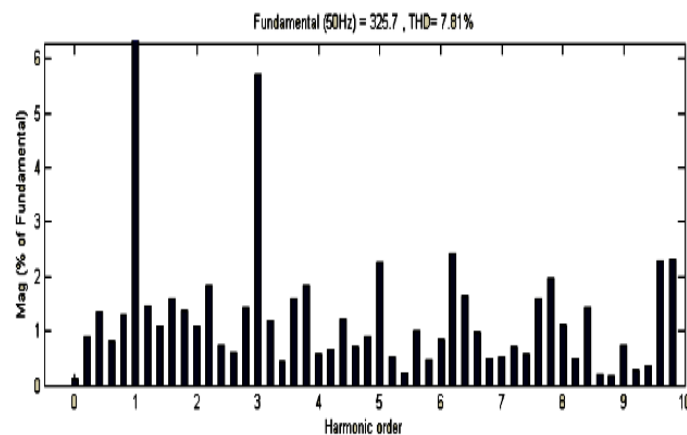


Fig.9 FFT spectrum of load voltage of MLI ($F_{opt} = 3950\text{Hz}$)

The FFT spectrum for the load current waveform is shown in Fig.10

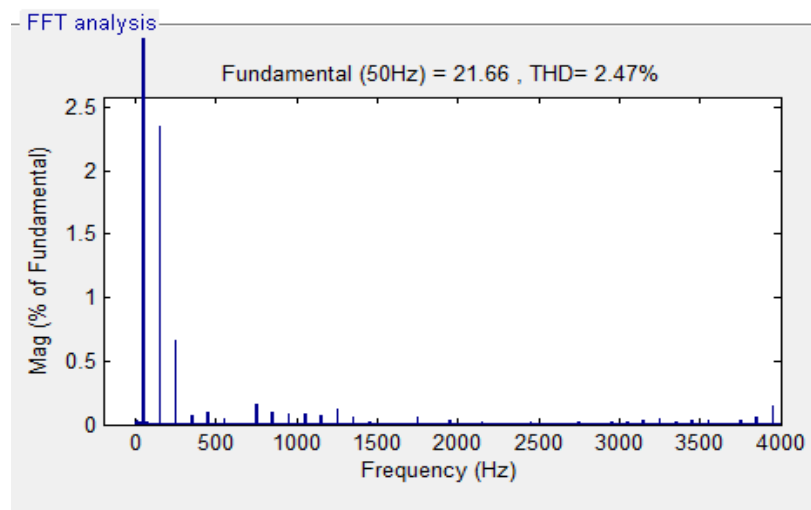


Fig.10 FFT Spectrum for the load current of MLI

The fundamental component of voltage [18-19] increases with increase in switching frequency and is higher for inverted sine carrier compared to the conventional triangular carrier which is shown in Fig.11.

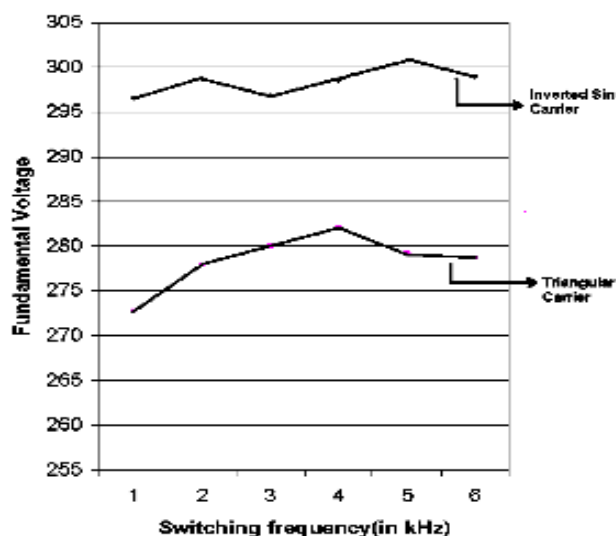


Fig.11 Fundamental voltage Vs Switching frequency

From the above results, it is observed that ISPWM gives an enhanced fundamental voltage and reduced total harmonic distortion.

V. EXPERIMENTAL RESULTS

To experimentally validate the asymmetric cascaded H-bridge multilevel inverter using the proposed modulation, a prototype of three-phase seven-level inverter has been built using FSBB20CH60 smart power module (SMP) as the switching devices shown in Fig.1. The SMP uses IGBT as the power device and it provides optimized circuit protection and drive matched to low loss IGBT. Hardware can be divided into the following sections: a Xilinx Spartan board, 7-level MLI setup (optocoupler, driver & power circuit) and load configuration. Xilinx Spartan 3-A DSP trainer is employed to generate the pulses required to trigger the IGBTs. The PWM pins AE3 to AF8 are used to generate the gating pulses of the respective devices. Transformers, Bridge Rectifiers and Voltage Regulators are used to form 12V DC supply for opto-couplers. Optocoupler consists of LED and a phototransistor. When an electrical signal is applied to the input of the opto-isolator, its LED lights, its light sensor then activates, and a corresponding electrical signal is generated at the output. The experimental setup for the asymmetric MLI is shown in Fig.12.

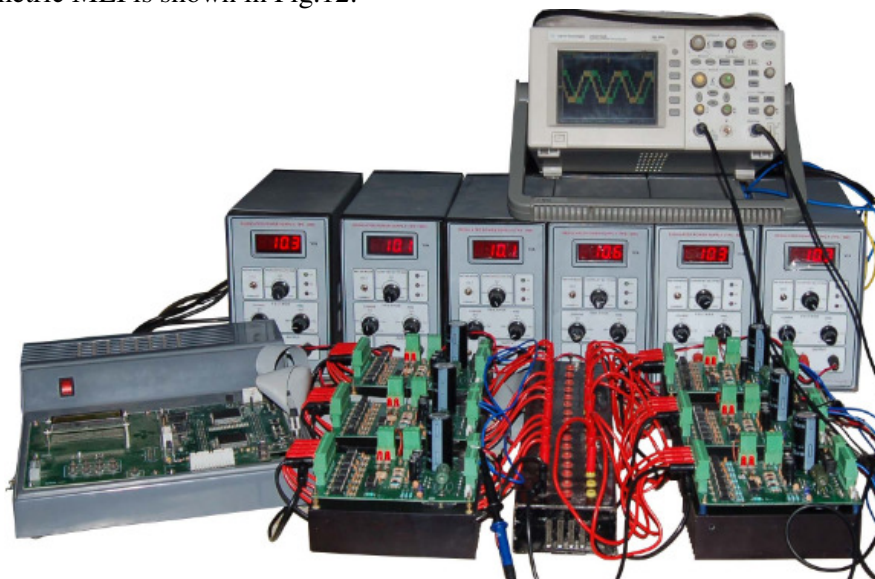


Fig.12 Photograph of the prototype of 3-phase asymmetric MLI

The experimental phase voltage waveforms is shown in Fig.13

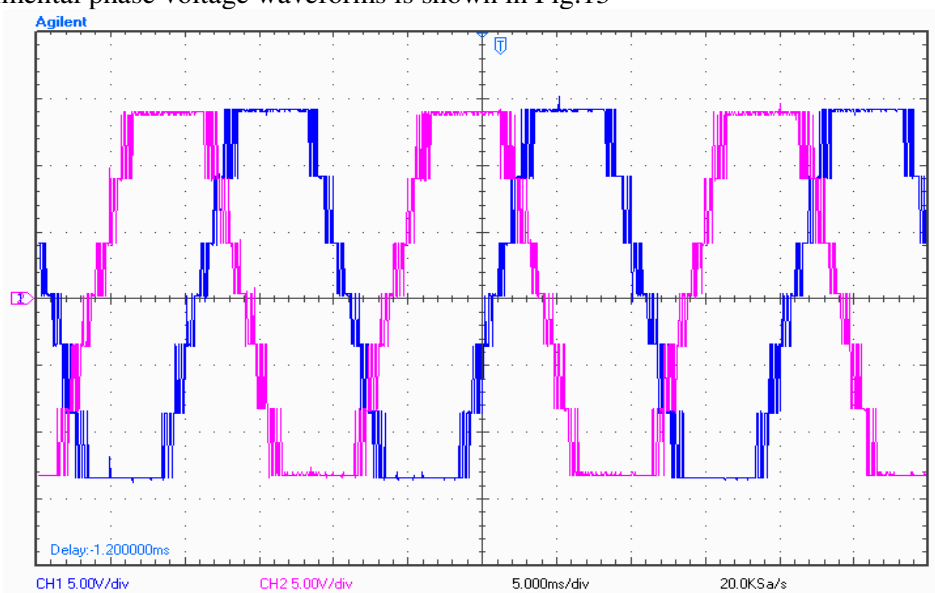


Fig.13 Experimental phase voltage waveform for 3-phase asymmetric cascaded MLI

The experimental line –line voltage waveform for MLI employing ISPWM technique is shown in Fig.14.

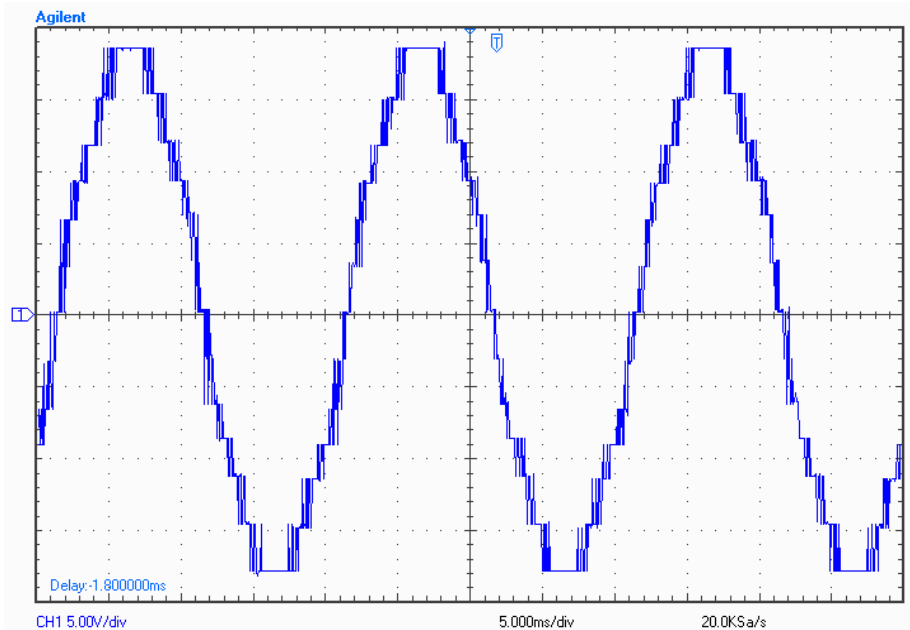


Fig.14 Experimental phase voltage waveform for 3-phase asymmetric cascaded MLI

VI. CONCLUSIONS

A three-phase asymmetric cascaded multilevel inverter employing inverted sine PWM technique has been investigated. From the simulation and experimental results, it is observed that ISPWM technique provides an enhanced fundamental voltage and reduced THD compared to the conventional triangular based PWM. Line-line voltage waveform of the proposed MLI topology produces 13-level which greatly reduces the harmonics in the output waveform. Asymmetric MLI Topology uses reduced number of DC sources thus decreasing the complexity and the cost of the circuit. This can be important in the high power quality cascaded multilevel inverters which require several voltage sources and knowledge of the DC voltage levels. By increasing the number of steps, waveform approaches the desired sinusoidal shape and THD is reduced. The proposed modulation strategy can be applied to multilevel inverter powered by fuel cells in place of DC sources and it has a greater scope in applications involving electrical vehicles and distributed generation.

REFERENCES

- [1]. Fang Zheng Peng, Jih-Sheng Lai, and Rodriguez, J. "Multilevel inverters: a survey of topologies, controls, and applications", *Industrial Electronics, IEEE Transactions*, Vol. 49, issue:4, pp. 724-738, Aug 2002.
- [2]. B.L.Mathur and R.Seyezhai "Harmonic Evaluation of Multicarrier PWM Techniques for Cascaded Multilevel Inverter". *International Conference on Electrical Engineering Applications ICEEA 2008*, Algeria.
- [3]. M.Calais, L. J. Borle and V.G. Agelidis, "Analysis of Multicarrier PWM Methods for a Single-phase Five Level Inverter", in *Proc. 32nd IEEE Power Electronics Specialists Conference, PESC'01*, July 2001, pp 1351-1356.
- [4]. Radan, A.H.Shahirinia, and M Falahi, "Evaluation of Carrier-Based PWM Methods for Multi-level Inverters", in *Proc. IEEE International Symposium on Industrial Electronics, ISIE 2007*, pp. 389-394, 2007.
- [5]. S. Mariethoz, and A.C. Rufer, "Design and control of asymmetrical multi-level inverters," in *Proc. IEEE 28th Annual Conference, IECON 02*, Vol. 1, pp. 840-845, 2002.
- [6]. D. Zhong, , L.M Tolbert, J.N. Chaiisson, and B.Ozpineci, "A Cascaded Multilevel Inverter Using a Single DC Source", in *Proc. Applied Power Electronics Conference and Exposition*, Vol. 2, pp. 426-430, 2006.

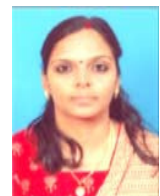
- [7]. P.Dananjayan, S.Jeevananthan, R.Nandhakumar “Inverted Sine Carrier for Fundamental Fortification in PWM Inverters and FPGA Based Implementations”. Serbian Journal of Electrical Engineering, Vol. 4, No. 2, pp. 171-187, November 2007.
- [8]. R. Seyezhai and B.L.Mathur, “Implementation and control of Variable Frequency ISPWM Method for an Asymmetric Multilevel Inverter”, European Journal of Scientific Research, Vol. 39, Issue 4, 2010, pp. 558-568.
- [9]. R. Seyezhai and B.L.Mathur, “Performance Evaluation of Inverted Sine PWM Techniques for an Asymmetric Multilevel Inverter”, Journal of Theoretical and Applied Information Technology, Vol. 2, No. 2, 2009, pp. 91-98.
- [10]. M.G.H. Aghdam, , S.H. Fathi, and B. Gharehpetian, “A Novel Carrier based PWM Method for THD reduction in Asymmetric Multilevel Inverter”, in Proc. International Conference on Renewable Energies and Power Quality, 2008, pp. 105-110.
- [11]. Z.D.Far, , A. Radan, and M.D. Far, “Introduction and Evaluation of novel multi-level carrier based PWM strategies using a generalized algorithm”, in Proc. European Conference on Power Electronics and Applications, EPE’07, pp. 1-10, 2007.
- [12]. L.M Tolbert, and T.G. Habetler, “Novel multilevel inverter Carrier based PWM Methods”, in Proc. IEEE IAS 1998 Annual Meeting, St.Louis, Missouri, pp. 1424-1431, 1998.
- [13]. S. Yuvarajan and A. Khoei, “A Novel PWM for Harmonic Reduction and its Application to AC-DC Converters”, IETE Journal of Research, Vol. 48, No. 2, pp. 85-92, 2002.
- [14]. S. Yuvarajan and Abdullah Khoei, “A novel PWM for harmonic reduction and its application to ac-dc converters,” IETE Journal of Research, Vol. 48, No. 2, March-April 2002.
- [15]. D. Quek and S. Yuvarajan, "A novel PWM scheme for harmonic reduction in power converters," Proc. of International conference on Power Electronics and Drive systems, Singapore, February 1995.
- [16]. N. A. Azli and Y.C. Choong, “Analysis on the performance of a Three- phase Cascaded H-Bridge Multilevel inverter”, in Proc. First International Conference on Power and Energy PECON 2006, Malaysia, pp. 405-410, 2006.
- [17]. M.G.H.Aghdam, S.H.Fathi and B.Gharehpetian, “Analysis of multicarrier PWM methods for asymmetric multilevel inverter”, in Proc. 3rd IEEE Conference on Industrial Electronics and Applications, ICIEA’08, pp. 2057- 2062, 2008.
- [18]. H.Keivani, M.R.Askari, F.Kavahnia, Aghdam, A. Mohammadi, "Novel multicarrier PWM method for a three- phase cascaded H-bridge multilevel inverter", in Proc. 41st International Universities Power Engineering Conference, UPEC 2006, 6- 8 September 2006, Vol .2, pp- 593 – 597.
- [19]. Radan, A. H. Shahirinia, M. Falahi, “Evaluation of Carrier- Based PWM Methods for Multilevel Inverters” in Proc. IEEE International Symposium on Industrial Electronics, ISIE07, June 2007, pp.389-394.

ACKNOWLEDGEMENT

The authors wish to express her gratitude to the management of SSN Institutions, Chennai, India and Vi Microsystems, Chennai for providing the laboratory and computational facilities to carry out this work.

AUTHOR PROFILE

R. Seyezhai obtained her B.E. (Electronics & Communication Engineering) from Noorul Islam College of Engineering, Nagercoil in 1996 and her M.E in Power Electronics & Drives from Shanmugha College of Engineering, Thanjavur in 1998 and PhD from Anna University, Chennai. She has been working in the teaching field for about 13 Years. She has published several papers in International Journals and International Conferences in the area of Power Electronics & Drives. Her areas of interest include SiC Power Devices, Multilevel Inverters, Modeling of fuel cells, Design of Interleaved Boost Converter, Multiport DC-DC Converter and control techniques for DC-DC Converter.



ON THE SUNSPOT TIME SERIES PREDICTION USING JORDON ELMAN ARTIFICIAL NEURAL NETWORK (ANN)

Rohit R. Deshpande and Athar Ravish Khan

Department of Electronics and telecommunication Engg.,
Jawaharlal Darda Institute of Engg. & Tech Yavatmal Maharashtra State, India.

ABSTRACT

In this paper, multi step ahead prediction of monthly sunspot real time series are carried out. This series is highly chaotic in nature [7]. This paper compares performance of proposed Jordan Elman Neural Network with TLRNN (Time lag recurrent neural network), and RNN (Recurrent neural network) for multi-step ahead (1, 6, 12, 18, 24) predictions. It is seen that the proposed neural network model clearly outperforms TLRNN(Time lag recurrent neural network), and RNN(Recurrent neural network) in various performance measures such as MSE (Mean square error), NMSE (Normalized mean square error) and r (correlation coefficient) on testing as well as training data set for 1,6,12,18,and 24 months ahead prediction of sunspot time series. Parameters are calculated by using software, "Neurosolution 5.0". Neurosolution is an object oriented environment for designing, prototyping, simulating, and deploying artificial neural network (ANN) solutions [26].

KEYWORDS: MSE (Mean square error), NMSE (Normalized mean square error), r (Correlation coefficient), k (Number of step ahead)

I. INTRODUCTION

The main motivation for analysis and research of Sunspot time series is to predict the future and understand the fundamental feature and effects on nature and human life. Artificial neural network is used for solving the real world problem [6]. This is mainly due to their ability to deal with nonlinearities, non-stationary and non gaussianity [24]. The modelling and analysis of chaotic time series has also attracted the attention of many researchers [4]. In this paper, Jordan Elman NN is found an optimal NN as compared to TLRNN & RNN for Sunspot time series prediction. Jordan Elman Neural Network is trained for multi step ahead prediction and the results are compared with reference to the MSE (mean square error), NMSE (normalized mean square error), and r (correlation coefficient), on testing as well as training data sets for short prediction. The number of experiments is carried out by changing various parameters like Error criterion (L1, L2, L3, L4, L5, and L_{∞}), number of iterations, learning rule, percentage of training & testing data sets and transfer functions [1]. Thus the optimum neural network model is proposed for short time prediction of Sunspot time series. Previously, some work has been done on the prediction of sunspot time series. It has been shown that the Fully Recurrent Neural Network (FRNN) model gives the good results for prediction [1]. In this paper; we have compared the various parameters for prediction and proposed the Jordan Elman neural network, which gives perhaps the better results. The remainder of the paper is organized as follows. In the next section we present the effective use of Artificial Neural Network for prediction. Section 3 deals with the introduction of proposed Jordan Elman neural network. Section 4 deals with the various performance measures on the basis of which the prediction has been carried out. Section 5 deals with the introduction of Neurosolution, which is used as a tool for the analysis. The graph representing calculation of number of processing elements, parameters obtained for various networks

in tabular form, the relation between actual and desired output along with the scatter plots for the variation of number of epochs has been shown in section 6. The paper ends with the comparison tables showing results for various networks and conclusion.

II. NEURAL NETWORKS FOR PREDICTION

The beginning of the prediction of the time series was made in the past century with the introduction of the autoregressive model for the prediction of the annual number of sunspots. Time series are samples of system's behaviour over discrete time values. The neural networks ability to cope up with the nonlinearities, the speed of computation, the learning capacity and the accuracy, made them valuable tools of prediction. To predict the evolution over time of a process implies the prediction of the future values of time series describing the process [1], [4]. Time series prediction with classical methods relies on some steps to be followed in order to perform an analysis of the time series, including modelling, the identification of the model and finally parameter estimation. The most difficult systems to predict are a) those with insufficient amount of data in the data series (for instance chaotic signals); b) the systems whose data series are obtained through observation and measurements, in this case data being possibly affected by measurement errors; c) systems whose behaviour varies in time [7]. The artificial neuron for prediction receives the delayed signal at its input $y(k-i)$, i is the number of inputs, the inputs are transmitted through $i+1$ multipliers for scaling, then the scaled inputs are linearly combined to form an output signal that is also passed through an activation function in order to obtain an output signal. The model of the neuron is schematized as shown in figure 1, [2].

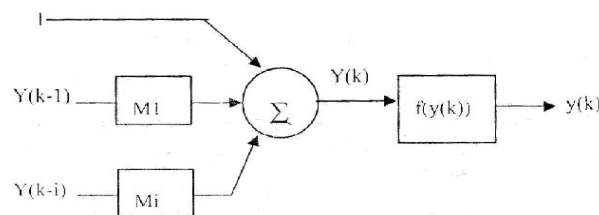


Figure1. The artificial neuron for prediction ($M1, \dots, i$ are multipliers)

Consider $x(t)$, the time series data collected in order to build a model. The data in the data series are samples of $x(t)$ with a time step k . The prediction of the time series values can be made single step ahead, in this case one looking for a good estimation $X1(t+1)$ of $X(t+1)$ or multi step ahead prediction, in this case looking for a good estimate $X1(t+nk)$ of $X(t+nk)$, n being the number of steps ahead [9]. The first and most common method for the prediction of time series consist in using M past values or M -tuples as inputs and one output.

III. JORDAN ELMAN NN

The theory of neural networks with context units can be analyzed mathematically only for the case of linear PEs. In this case the context unit is nothing but a very simple low pass filter. A low pass filter creates an output that is a weighted (average) value of some of its more recent past inputs. In the case of the Jordan context unit, the output is obtained by summing the past values multiplied by the scalar as shown in the figure 2. Notice that an impulse event $x(n)$ (i.e. $x(0)=1$, $x(n)=0$ for $n>0$) that appears at Time $n=0$, will disappear at $n=1$. However, the output of the context unit is $t1$ at $n=1$, $t2$ at $n=2$, etc. This is the reason these context units are called memory units, because they "remember" past events. t should be less than 1, otherwise the context unit response gets progressively larger (unstable). The Jordan network and the Elman network combine past values of the context units with the present inputs to obtain the present net output. The input to the context unit is copied from the network layer, but the outputs of the context unit are incorporated in the net through adaptive weights [9], [24]. NeuroSolutions uses straight back propagation to adapt all the network weights [2].

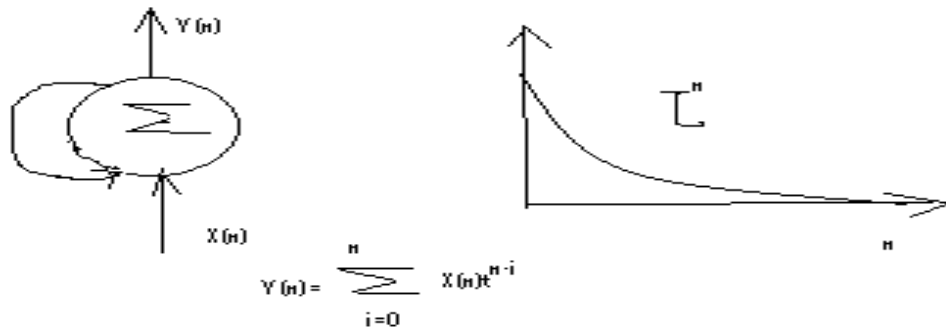


Figure 2. Context unit response

In the Neural Builder, the context unit time constant is pre-selected by the user. One issue in these nets is that the weighting over time is kind of inflexible since we can only control the time constant (i.e. the exponential decay). Moreover, a small change in t is reflected in a large change in the weighting (due to the exponential relationship between time constant and amplitude). In general, we do not know how large the memory depth should be, so this makes the choice of t problematic, without a mechanism to adapt it. See time lagged recurrent nets for alternative neural models that have adaptive memory depth. The Neural Wizard provides four choices for the source of the feedback to the context units (the input, the 1st hidden layer, the 2nd hidden layer, or the output). In linear systems the use of the past of the input signal creates what is called the moving average (MA) models. They represent well signals that have a spectrum with sharp valleys and broad peaks. The use of the past of the output creates what is called the autoregressive (AR) models. These models represent well signals that have broad valleys and sharp spectral peaks. In the case of nonlinear systems, such as neural nets, these two topologies become nonlinear (NMA and NAR respectively). The Jordan net is a restricted case of an NAR model, while the configuration with context units fed by the input layer is a restricted case of NMA. Elman's net does not have a counterpart in linear system theory. As you probably could gather from this simple discussion, the supported topologies have different processing power, but the question of which one performs best for a given problem is left to experimentation[3],[7].

IV. PERFORMANCE MEASURES

The generalization performance of the network is validated on the basis of the following parameter.

4.1. MSE (Mean Square Error)

It is the average of the square of the difference between each output processing element and the desired output. It is used to determine how well the network output fits the desired output, but it doesn't reflect whether two sets of the data move in same direction [1].

Y_{ij} = network output for exemplar i at PE j
 P = number of output PEs (processing elements)
 N = number of exemplar in datasheet
 D_{ij} = desired output for exemplar i at PE j

4.2. NMSE (Normalized Mean square error)

$$NMSE = \frac{p \cdot n \cdot MSE}{\sum_{j=0}^p \left[\frac{N \sum_{i=0}^N d_{ij}^2 - \left(\sum_{i=0}^N d_{ij} \right)^2}{N} \right]} \quad (1)$$

NMSE is given as -

P = number of output PEs (processing elements)
 N = number of exemplar in datasheet
 dij = desired output for exemplar i at PEj

4.3 Correlation coefficient (r)

It reflects whether the two sets of data moves in same direction. The correlation coefficient is between the network output X and the desired output D is

$$r = \frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{\sqrt{\sum_i \frac{(d_i - \bar{d})^2}{N}} \sqrt{\sum_i \frac{(x_i - \bar{x})^2}{N}}} \quad (2)$$

Where,

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{and} \quad \bar{d} = \frac{1}{N} \sum_{i=1}^N d_i \quad (3)$$

The correlation coefficient is confined to the range [-1, 1]. When r = 1, there is perfect positive linear correlation between X & d that is they co vary means they vary by the same amount [T. Edwards, D.S.W Tansley, R.J. Frank, N. Davey .etal], [Lapedes A. and Farber R.etal]. The above parameters are calculated by using software, "Neurosolution 5.0"

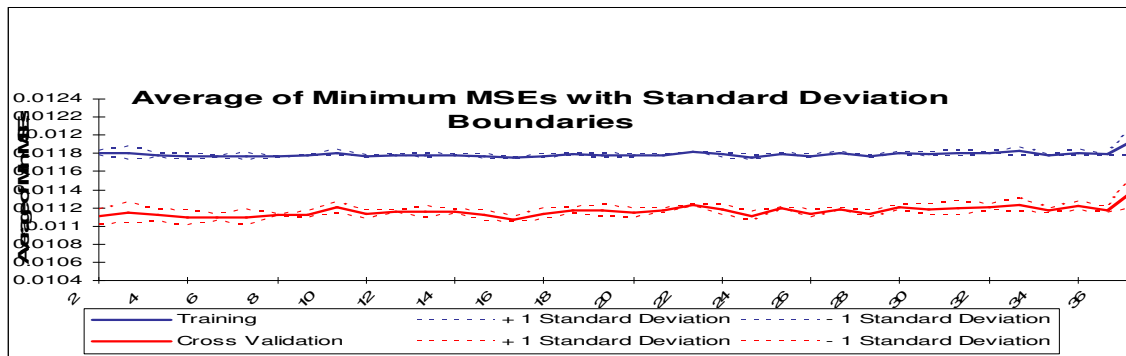
V. NEUROSOLUTION 5.0

Neurosolution is an object oriented environment for designing, prototyping, simulating, and deploying artificial neural network (ANN) solutions. This environment is inherently modular, as a neural network breaks down into a fundamental set of neural components. Individually simple, these components combine to form networks that can solve complex problems. Neurosolution supports a practically infinite number of neural models. It provides the user with an icon-based interface from which the neural components are arbitrarily interconnectable [2].

VI. CASE STUDY

The different neural network models like Jordan Elman, time lag recurrent network and recurrent neural network are trained for multi step ahead [1,6,12,18,24] predictions and the results are compared with reference to MSE (Mean square error), NMSE (Normalized Mean Square Error), r (Correlation coefficient) on testing as well as training data set for short term prediction. The number of experiments is carried out by changing various parameters like number of processing elements, number of hidden layers, number of iterations, transfer function, learning rule. Here we are finding out the number of processing elements for which the network gives minimum MSE.

Figure 3 shows the graph obtained by varying various parameters giving the number of processing elements for which the network gives minimum MSE. The Following Tables gives values of various parameters such as MSE (Mean square error), NMSE (Normalized Mean Square Error), and r (Correlation coefficient) for different look ahead values for different Neural Networks. Table 1 gives values of MSE, NMSE, and r of Jordan Elman NN for multi step ahead (1, 6, 12, 18, 24) prediction for Sunspot time series.



Best Networks	Training	Cross Validation
Hidden 1 PEs	15	5
Run #	3	2
Epoch #	1000	1000
Minimum MSE	0.011739974	0.011022748
Final MSE	0.011739974	0.011022748

Figure 3. Graph and Table of varying a parameter

Similarly table 2 and table 3 shows various values of MSE, NMSE, and r for Time Lagged Recurrent Neural Network and Recurrent Neural Network. Comparing the performance measures like MSE, NMSE and r for 60% samples as training, 25% samples as testing and 15% as cross validation, it is found that the performance of the selected model is optimal for 15 neurons in the hidden layer. Next the proposed model is trained with different error criterion L1, L2, L3, L4, L5 and L ∞ .

Table1. Type of network-Jordan Elman NN

For South				For North		
K (step)	MSE	NMSE	R	MSE	NMSE	r
1	0.00301352	0.12175413	0.9372644	0.00039479	0.01198851	0.99412255
6	0.00315088	0.12097912	0.9384633	0.00131084	0.00339036	0.98179276
12	0.01120112	0.23450117	0.8901836	0.01142852	0.16956471	0.93439364
18	0.00992762	0.43738162	0.8810052	0.01900812	0.26451196	0.91967542
24	0.00969921	0.47540716	0.8735284	0.01007496	0.27224255	0.90736715

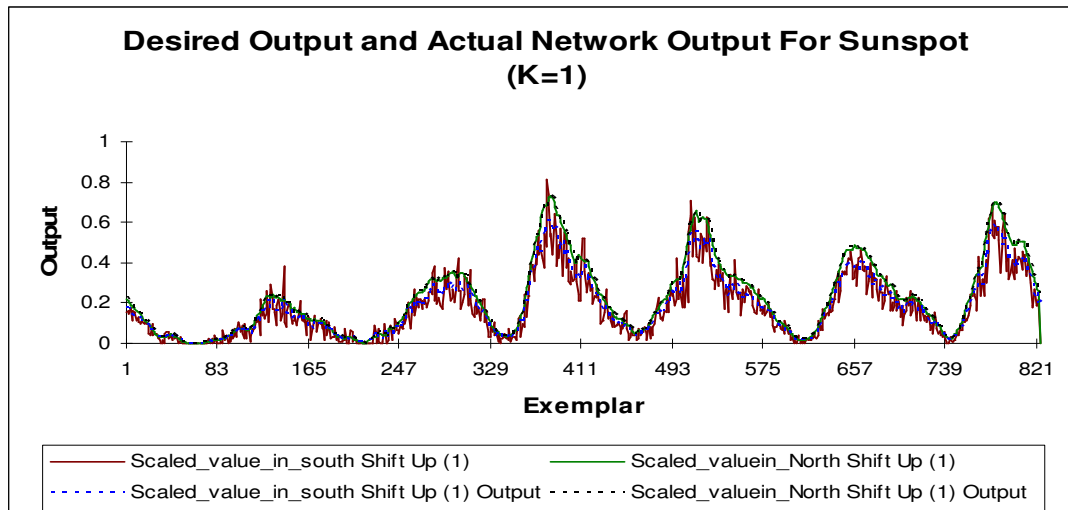
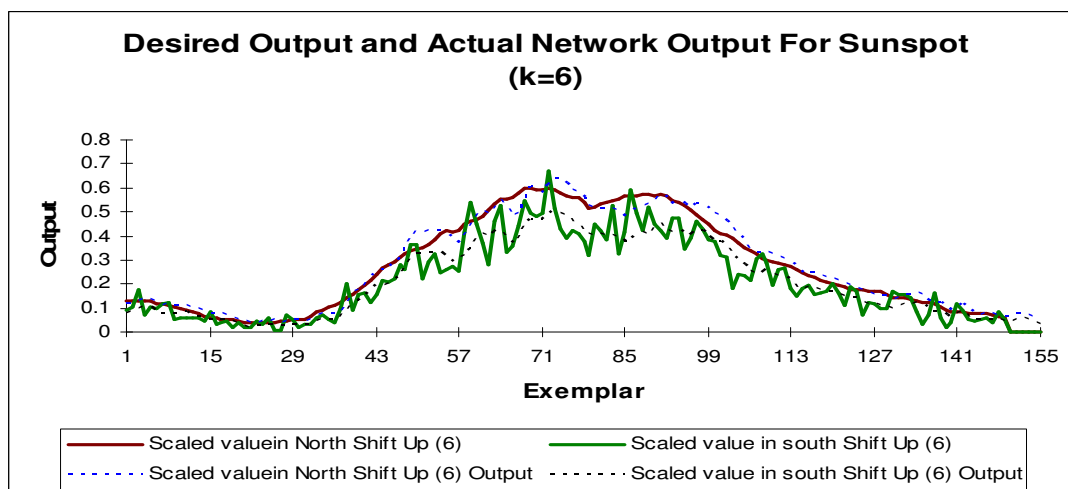
Table 2. Type of Network Time Lagged Recurrent Neural Network

For South				For North		
K (step)	MSE	NMSE	R	MSE	NMSE	r
1	0.0031452	0.1338280	0.930992	0.001002	0.030435	0.9832722
6	0.0038522	0.1372878	0.927177	0.001627	0.004266	0.9726262
12	0.0116022	0.2429242	0.882282	0.011626	0.164626	0.9302727
18	0.0165966	0.3459452	0.840728	0.021829	0.322258	0.8775592
24	0.0997141	0.3538383	0.813266	0.031833	0.363333	0.8610236

Table 3. Type of Network Recurrent Neural Network

For South				For North		
K(step)	MSE	NMSE	R	MSE	NMSE	r
1	0.0023245	0.1339322	0.951992	0.004242	0.014249	0.9812222
6	0.0390393	0.1463663	0.927135	0.001634	0.042527	0.9788397
12	0.0143535	0.2967211	0.884222	0.014892	0.214893	0.9236422
18	0.0178532	0.3721266	0.819626	0.021780	0.321859	0.8610656
24	0.0128535	0.4238543	0.777512	0.013118	0.302349	0.8423390

The best combination network is then trained and tested for different transfer functions such as a) Tanh axon b) Sigmoid c) Linear Tanh axon d) Linear Sigmoid. The proposed Jordan Elman NN model is trained for the best combinations resulted for training and testing exemplars and it is experimented for 1000 to 20000 iterations for getting an optimum results for each multi step ahead ($k=1, 6, 12, 18, 24$) of chaotic sunspot time series .The number of epochs are varied from 1000 to 20000 in the step 2000 and the graphs are plotted for all the steps. These Graphs are shown in Figure 4 to Figure 8.

**Figure 4.** Graph For ($k=1$)**Figure 5.** Graph For ($k=6$)

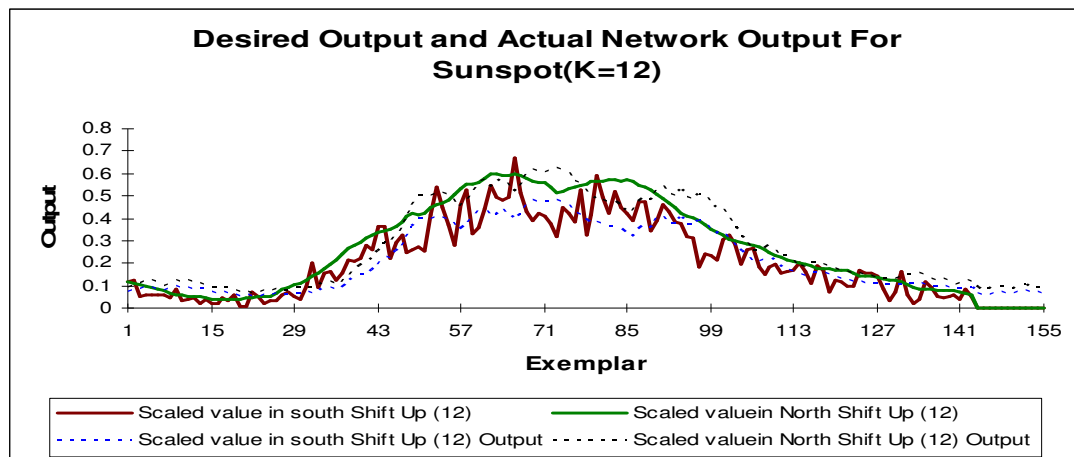


Figure 6. Graph For (K=12)

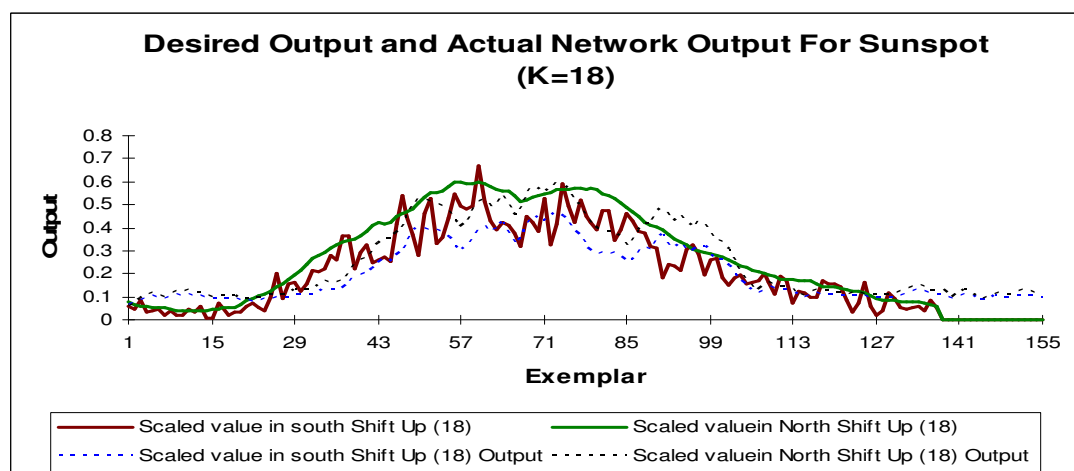


Figure 7. Graph For (K=18)

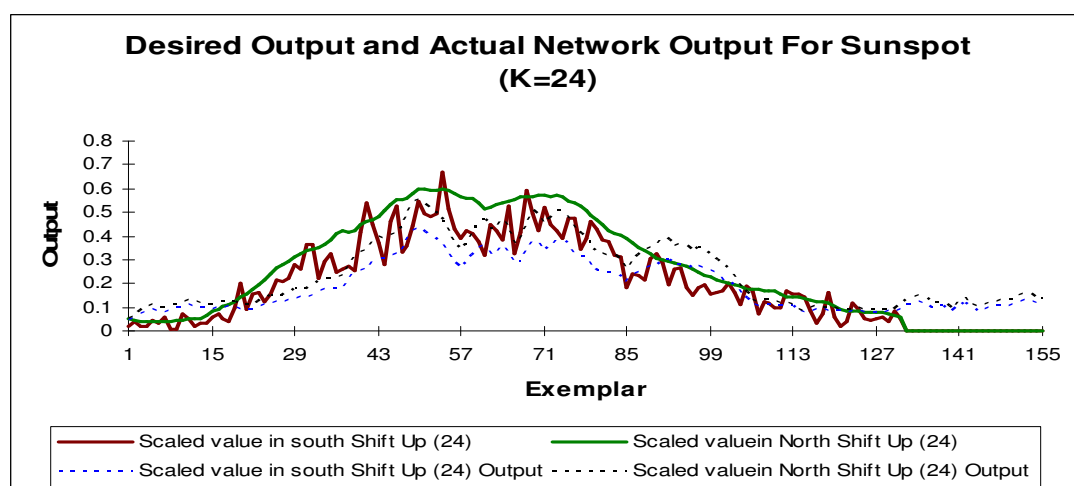


Figure 8. Graph For (K=24)

Then, we have varied number of epochs means number of iterations. Epochs are varied from 1k to 20 k (1000 to 20000). By analyzing its results we have plotted the scatter plots of number of epochs and variation of MSE, NMSE, & r.

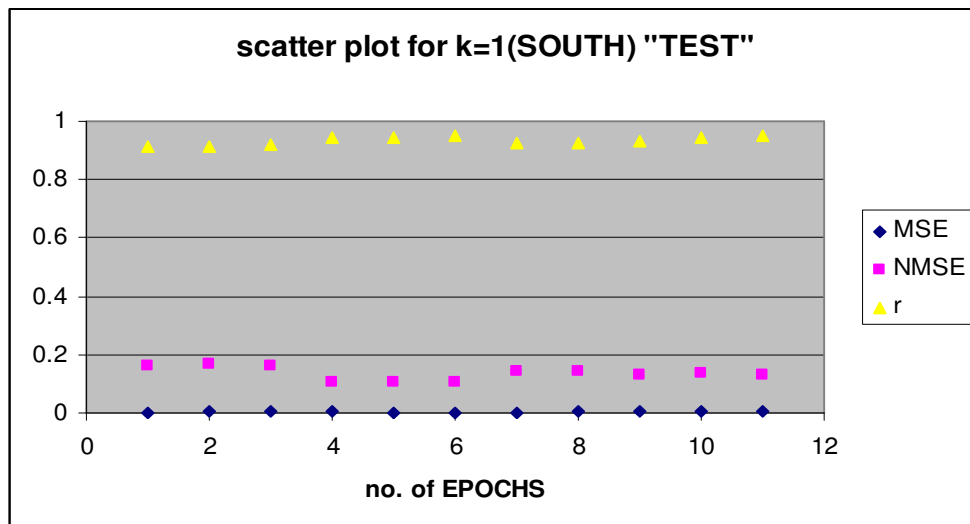


Figure 9. Scatter Plot For K=1

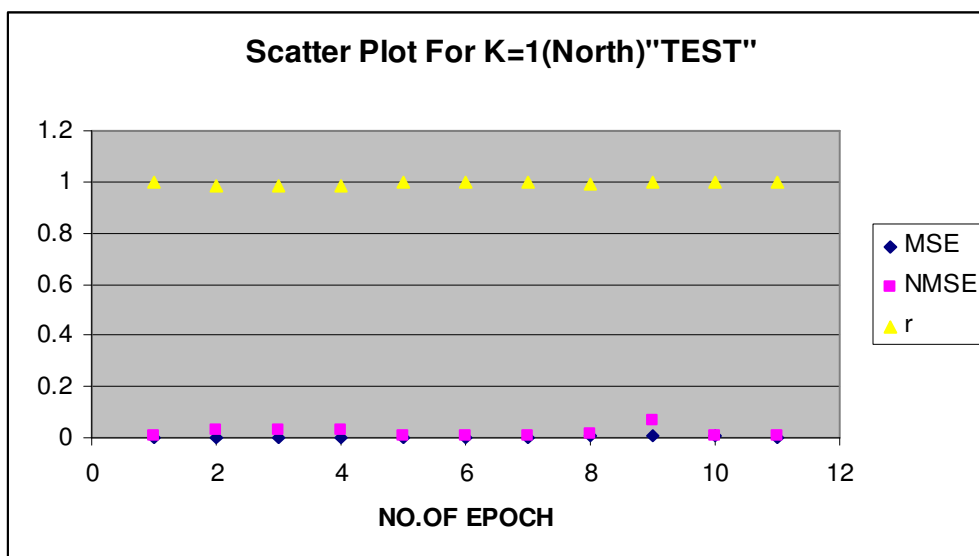


Figure 10. Scatter Plot For K=1

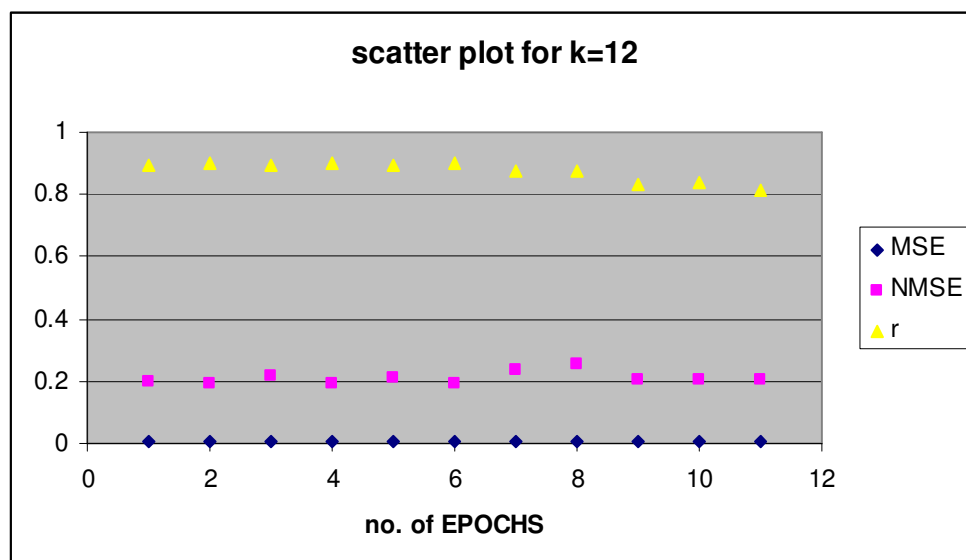


Figure 11. Scatter Plot For K=12

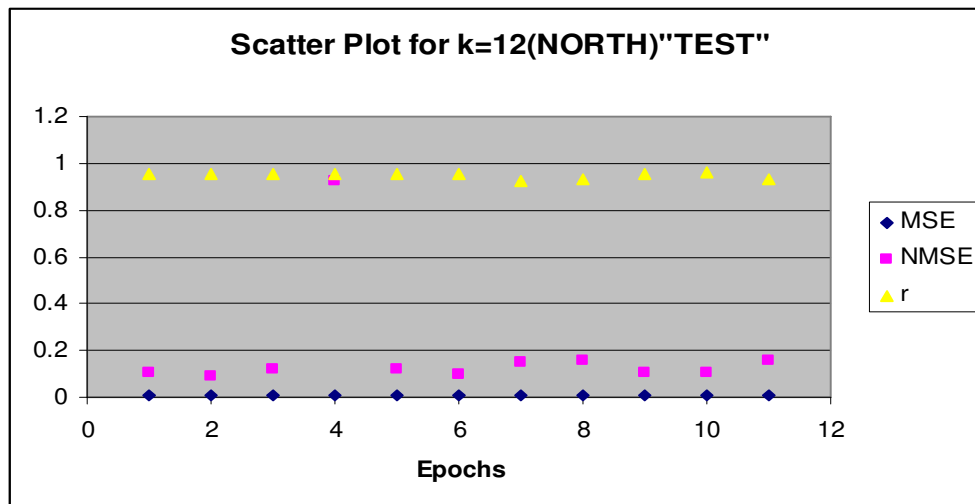


Figure 12. Scatter Plot For K=12

VII. RESULTS

Comparison Tables of JENN, TLRNN, FRNN For chaotic time series (k= 1 & k=6)

Table 4. For k=1

S. No.	NN MODEL	SOUTH			NORTH		
		R	MSE	NMSE	r	MSE	NMSE
1	JORD-ELMN	0.9398	0.0028	0.1170	0.9931	0.0004	0.0138
2	TLRNN	0.9394	0.0029	0.1181	0.9853	0.0010	0.0304
3	FRNN	0.9155	0.0040	0.1618	0.9812	0.0012	0.0372

Table 5. For k=6

S. No.	NN MODEL	SOUTH			NORTH		
		r	MSE	NMSE	r	MSE	NMSE
1	JORD-ELMN	0.9384	0.0031	0.1209	0.9817	0.0013	0.0339
2	TLRNN	0.9292	0.0038	0.1398	0.9809	0.0014	0.0380
3	FRNN	0.9271	0.0390	0.1461	0.9788	0.0016	0.0372

Tables 4 and 5 shows comparison between Jordn-Elman NN, TLRNN, and FRNN for shift(k=1,6). Thus here we have compared various NN on the basis of MSE, NMSE, and r, and an optimal NN for prediction of Sunspots time series is found.

VIII. CONCLUSIONS

It is observed that Jordan Elman Network is able to predict Sunspots time series quite well in comparison with other networks. It is seen that MSE, NMSE & r of the proposed dynamic model for testing data set as well as training data set are significant than other neural models. The network is analyzed for different step ahead (k=1, 6, 12, 18, 24). The better results can be obtained for maximum

24 step ahead by using this network. The following parameters are obtained for 24 step ahead; MSE=0.009457611, NMSE=0.309630953, $r=0.87684609$ (for South), MSE=0.010051769, NMSE=0.281693745, $r=0.908057335$ (for North) For Epoch=16000. Here, we can see that the value of correlation coefficient is much closer to unity which makes the proposed network optimal one for prediction. In this paper, we have also studied the variation of error and correlation coefficient with the number of epochs. Also, the variation of actual and desired output can also be studied for different step ahead values (1, 6, 12, 18, and 24) by plotting the graphs. It can be seen that the proposed neural network closely follows the actual output.

REFERENCES

- [1] Sanjay L. Badjate, Sanjay M. Gulhane "Multi-Step Ahead Prediction Of Chaotic Time Series Using Recurrent Neural Network Models ". International Journal Of Computer Science And Applications Vol. 4, No. 1, April / May 2011 ISSN: 0974-1003
- [2] Alluada Neurosolution 5.0.
- [3] J. Hertz, A. Krogh, R.G. Palmer. "Introduction to theory of neural computation", vol-1-Addison-Hesley 1991.
- [4] William Ditto, Toshinori Munakata "Principles and applications of chaotic system".
- [5] Sorin Vlad "On the prediction methods using neural networks".
- [6] Tuevo kohonen "self organizing feature map", third edition, Springer 2001, p.28
- [7] K.T. Alligood, T.D. Sauer, J.A. Yorke, "CHAOS-An introduction to neural network"
- [8] D. Hammerstram, "Working with neural network", IEEE spectrum, 1993, pp.46-53.
- [9] J.M. Zurada "Introduction to Neural Systems", Jaico Publication, 2004.
- [10] Yachanan shachmarove: "Applying neural network to business, economics"
- [11] Haykin S. Adaptive Filter Theory. Prentice-Hall, Englewood Cliffs, 1991.
- [12] Haykin S. Neural Networks, a Comprehensive Foundation. MacMillan, 1994.
- [13] Hebb D. The Organization of Behavior. Wiley, New York, 1949.
- [14] Hecht-Nielsen R. NeuroComputing. Addison Wesley, 1990.
- [15] Hertz J., Krogh A. and Palmer R. Introduction to the Theory of Neural Networks. Addison Wesley, 1991.
- [16] Hopfield J. "Neural networks and physical systems with emergent collective computational abilities." Proc. Natl. Acad. Sci. (USA) 79, 2554-2558, 1982.
- [17] Jacobs R. "Increased rates of convergence through learning rate adaptation." Neural Networks 1, 295-307, 1988.
- [18] Jordan M. "Attractor dynamics and parallelism in a connectionist sequential machine." Proc. 8th Annual Conf. on Cognition Science Society, pp 531-546, 1986.
- [19] Kohonen T. Self Organization and Associative Memory. Springer Verlag, 1988.
- [20] Kohonen T. "The self-organizing map." Proc. IEEE 78, 1464-1480, 1990.
- [21] Kosko B. Neural Networks and Fuzzy Systems. Prentice Hall, 1992.
- [22] Kung S.Y. Digital Neural Networks. Prentice Hall, 1993.
- [23] Lang K., Waibel A. and Hinton G. "A time delay neural network architecture for isolated word recognition." Neural Networks 3(1), 23-44, 1990.
- [24] Lapedes A. and Farber R. "Nonlinear signal processing using neural networks: prediction, and system modelling." LA-VR-87-2662, Los Alamos, 1987.
- [25] LeCun Y., Denker J., Henderson D., Howard R., Hubbard W. and Jackel L. "Handwritten digit recognition with a back propagation network." In Advances in Neural Information Processing Systems 2, (ed. Touretsky), pp 396-404, 1990.

- [26] Lefebvre C. and Principe J. "Object-oriented artificial neural network implementations." World Cong. Neural Networks 4, 436-439, 1993.
- [27] Lefebvre C. An Object-Oriented Methodology for the Analysis of Artificial Neural Networks. Master's Thesis, University of Florida, 1992.
- [28] Lippman R. "An introduction to computing with neural nets." IEEE Trans. ASSP Magazine 4,4-2, 1987.
- [29] Little W. and Shaw G. "Analytical study of the memory storage capacity of a neural network." Mathematical Biosciences 39, 281-290, 1978.
- [30] Makhoul J. "Pattern recognition properties of neural networks." Proc. 1991 IEEE Workshop on Neural Networks for Signal Processing, pp 173-187, 1992.
- [31] Marr D. "A theory of cerebellar cortex." Journal of Physiology 202, 437-470, 1969.
- [32] McCulloch W. and Pitts W. "A logical calculus of the ideas imminent in the nervous activity." Bulletin of Mathematical Biophysics 5, 115-133, 1943.
- [33] Minsky M. and Papert S. Perceptrons. MIT Press, 1969.
- [34] NeXTStep Operating System, NeXT Computer Documentation, 1991.

Authors

Rohit Ravindra Deshpande was born in Maharashtra, INDIA, in 1986. He received the B.E. degree in electronics and telecommunication engineering, from SGBA University Amravati Maharashtra INDIA, in 2008 and Pursuing M.Tech degree in Electronics Engg from RTMN University Nagpur. In 2008, he joined J.D.I.E.T. Yavatmal Maharashtra INDIA and worked as a lecturer. In June 2011 he became an Assistant Professor there. His current research interests include Artificial Neural Networks and time series prediction.



Athar Ravish Khan was born in Maharashtra, INDIA, in 1979. He received the B.E. degree in electronics and telecommunication engineering, M.E. degree in digital electronics from SGBA University Amravati Maharashtra INDIA, in 2000 and 2011 respectively. In 2000, he joined B.N College of Engineering Pusad and worked as lecturer. In 2006 he joined as lecturer in J.D Institute of Engineering and Technology Yavatmal, Maharashtra INDIA and in March 2011 he became an honorary Assistant Professor there. He is pursuing Ph.D. degree under the supervision of Prof. Dr. Sanjay M. Gulhane. His current research interests include digital signal processing, neural networks and wireless communications, with specific emphasis on UWB in underground Mines-Tunnels



SESSION DATA PROTECTION USING TREE-BASED DEPENDENCY

G. Shruthi¹, Jayadev Gyani², R. Lakshman Naik³, G. Sireesh Reddy⁴
^{1,2}Department of CSE, JITS, Warangal, A.P, India
^{3,4}Department of CSE, BITS, Warangal, A.P, India

ABSTRACT

Web applications have become very popular in nowadays. Web applications can be error prone and easily exploited by attackers, because of the implementation vulnerabilities. Securing web applications against implementation vulnerabilities is very important. Existing security solutions do not provide adequate support to protect web applications against broken session data dependencies in this paper we focus on one specific type of implementation vulnerability, namely broken dependencies on session data along with session data dependent vulnerabilities. This can lead to a variety of erroneous behavior at runtime. This paper shows how to guarantee the absence of runtime errors due to broken session data dependencies. The proposed solution provides the tree-based dependency to prove no-broken-data-dependencies property. A framework named PP4Wap (provable protection for web application) which will increase the reliability and security of data-centered web applications.

KEYWORDS: Security, Reliability, Data sharing, Web application, PP4Wap, Vulnerabilities, WAF.

I. INTRODUCTION

Now a day's almost of us use web application. Almost all of the services are provided via the World Wide Web (WWW) [1]. Without WWW we can't do anything easily. More and more organizations strongly depend on their correct functioning. Web application tends to be error prone and implementation vulnerabilities are readily and commonly exploited by attacks. So, to improve the reliability and security of web applications is most important in the field of software engineering and security engineering. This paper introduces a tree base dependency concept in session control. It is a protection approach against web application vulnerabilities related to session data dependencies. To prove this we have to take one specific implementation bug i.e. runtime errors due to broken data dependencies in data-centered applications. This is an application level vulnerability.

Current technologies such as anti-virus software programs and network firewalls provide the secure protection at the host and network levels, but not at the application level. Existing solution for the broken session data dependencies is Web Application Firewall (WAF) [2]. WAFs are applied to mitigate a range of vulnerabilities, including vulnerabilities to forceful browsing. But a malicious user will typically apply forceful browsing to exploit implementation-specific broken session dependencies in data-centered web applications in a more or less controlled way. So, existing security solutions do not provide adequate support to protect web applications against such implementation-specific bugs. Proposed solution combines development-time program annotation static verification and runtime checking and introduces tree-based dependency in session control to prove no-broken-data-dependency property.

The rest of this paper is structured as follows: Section 2 provides some background information on indirect data sharing and broken session data dependencies in data-centered Web applications. In Section 3, the presented work is related to existing research in program verification and Web security.

Section 4 defines the requirements for a solution. In Section 5 we present our solution to detect and prevent composition problems in data-centered Web applications. Section 6 defines the implementation of the tree-based dependency concept. Section 7, presents the result of the paper. Section 8 summarizes the conclusion of this paper.

II. BROKEN SESSION DATA DEPENDENCIES

This section provides more detailed information on how to break the session data in data-centered applications and why WAF fails to protect web applications against broken data dependencies [3].

2.1. Indirect data sharing

Indirect data sharing is adopted in several component models and APIs such as JavaServlet containers [4], Pluggable Authentication Modules framework [5] and Java Spaces in sun's Jinni [6]. Generally web-applications are server-side applications that are invoked by this Web Clients (browsers), typically using the Hyper-Text Transport Protocol (HTTP).

The HTTP is a stateless application-level request/response protocol and has been in use on the WWW since 1990[7]. Since the protocol is stateless, each request is processed independently without any knowledge of previous requests. To enable the typical user's session concept in web applications, the web application needs to add session management on top of the stateless HTTP layer. Different techniques exist to embed web request within a user session such as the use of cookies, URL rewriting or hidden form fields [8]. Out of these three methods cookies is a best method, because cookie functionality is built in to most browsers. Thus no special coding is required to ensure session ID information is embedded within the pages served to the client browser.

In the repository architectural style [9], a system consists of a central data structure and set of components interacting with the repository. The data-centered application contains this architectural style.

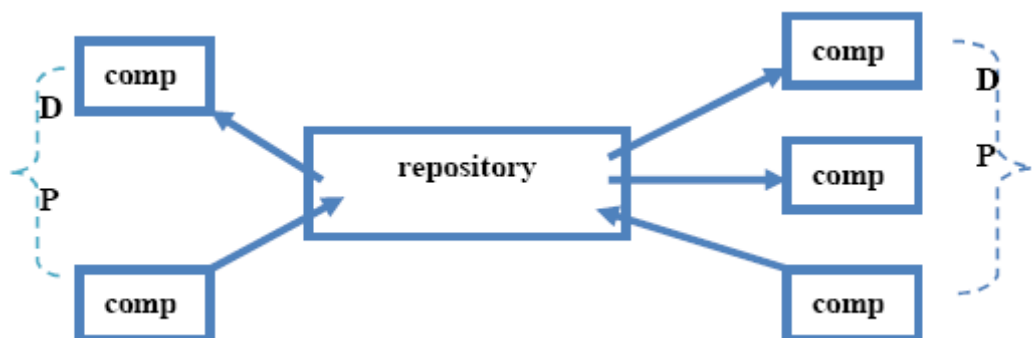


Fig. 1. Data dependencies (DP) in data-centered applications

In data-centered applications each component can indirectly interact with the other component, but all the components can directly interacting with the shared repository. The data centered applications are correctly composed with respect to the indirect data sharing if, at run time, each component is able to retrieve the data from the repository that it expects to find. Thus the correct functioning of a component depends on the run time state of the repository. At the time of execution, implicit semantically dependencies exist between the components and shared repository.

2.2. Broken data dependencies

Breaking data dependencies is a common risk in composing data-centered applications. We identified two common composition mismatches or exceptions in data centered applications. These are as follows:

2.2.1. NullPointerException

A data item is not available on the shared repository although a reading component expects it on the repository during execution.

2.2.2. ClassCastException

The type of an available data item does not correspond with the type expected by the reading component.

These mismatches lead to the runtime errors. Due to these runtime errors and loose coupling the data can break in data-centered applications. These broken data dependencies undermine the reliability and security of the web applications. Existing solution for this vulnerability is WAF, because this is an application level vulnerability.

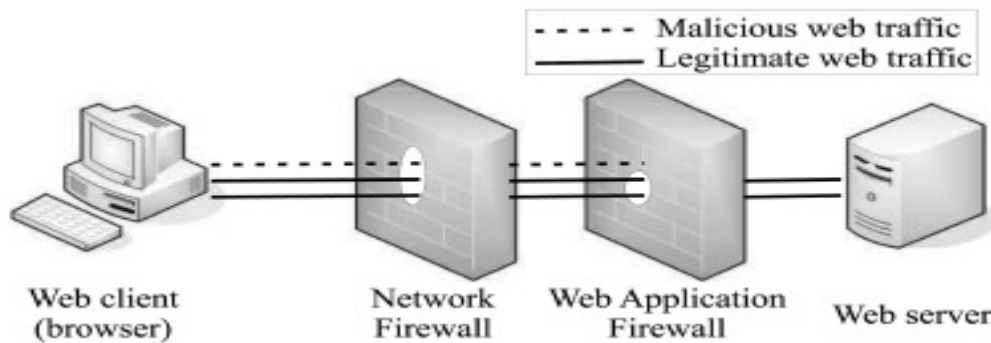


Fig. 2. Web Application Firewall infrastructure.

To counter web application vulnerabilities, Web Application Firewalls operate on the application layer and analyze web request between a browser and the web server [10]. Web application firewalls are applied to mitigate a range of vulnerabilities, including vulnerabilities to forceful browsing. By passing intended application flow in a web application can generally lead to unauthorized access to resource or unexpected application behavior [11]. Moreover, a malicious user will typically apply forceful browsing to break the session data dependencies in data-centered applications. Therefore existing security solutions do not provide adequate support to protect web applications. So, we need additional support for this. In proposed solution we have to take one composition property to reduce the runtime errors caused by broken data dependencies i.e. no-broken-data-dependencies [18].

III. RELATED WORK

The work presented in this paper is related to a broad spectrum of ongoing research. We only present some key pointers for each of the domains and, in more detail, for the domain most related to the proposed solution, namely, static and dynamic verification in Web application security.

Several implementation-centric security countermeasures for Web applications have already been proposed [12], [13], but most of them focus on injection attacks (SQL injection, command injection, XSS, etc.) and use tainting, pointer analysis, or data flow analysis. Our solution targets another set of implementation bugs, namely, bugs due to broken data dependencies on the server-side session state, and, to do so, we rely on the static and dynamic verification of component contracts. Gould et al. also aim to reduce the number of runtime errors in Web applications by applying static verification [14]. Their solution focuses on the reduction of SQL runtime exceptions and uses a static analysis tool to verify the correctness of all dynamically generated query strings within an application. Our solution is based on program annotations and we verify interactions between components and the non-persistent server-side state.

We combine in our solution static and dynamic verification to reduce the runtime enforcement overhead. The idea of combining static and dynamic verification is not new and has, for instance, already been adopted by Huang et al. in securing Web applications against Web vulnerabilities caused by insecure information flow, such as SQL injection, XSS, and command injection [15]. Their approach uses a lattice-based static analysis algorithm for verifying information flow based on type systems and the type state. The sections of the code considered vulnerable are automatically instrumented with runtime guards. In contrast, our approach aims to reduce runtime errors due to composition problems. In addition, our approach is based on program annotations and the verification of component preconditions.

In [16], generate bypass tests that check if an online Web application is vulnerable to forceful browsing or parameter tampering attacks. The bypass tests are black-box tests using data that circumvents client-side checks. They define three levels of fault injection: bypass tests at the value level, at the parameter level, and at the control flow level. Since the fault injections are based on violations of the client-side validations, they operate independently of the server implementation and do not give formal guarantees about the absence of bugs. In contrast, our verification approach is able to guarantee the absence of errors at the flow level and, in future work; we would like to investigate how well our approach is suited to counter errors at the other two levels as well.

Firewall configuration analysis is proposed to manage complex network infrastructures (such as networks with multiple network firewalls and network intrusion detection systems) [17]. These approaches aim to achieve efficiency and consistency between the different network layer security devices, whereas our approach focuses on the application-layer consistency between the WAF and the Web application.

For more than a decade, software architectures have been used to abstract reasoning about software systems from the source code level toward coarse-grained architectural elements and their interconnections [9], [19], [20], and [21]. Architectural styles abstract reoccurring patterns of components, connectors, and behavioral interactions within different software architectures and try to capture the advantages or main characteristics of a particular architectural style, as well as the constraints introduced by the style. In [9], Shaw and Garlan proposed taxonomy of different architectural styles, including the data-centered style.

To support architecture-based reasoning, (semi)formal modeling notations and analysis techniques are required. Several Architecture Description Languages (ADLs) are proposed for architectural specification and analysis. Although these ADLs strongly vary in the abstractions and analysis capabilities they provide, most ADLs explicitly provide abstractions for components, connectors, and their behavioral interactions, as well as tool support for analysis and architecture-based development [21]. However, in most cases, a discontinuity exists between the architectural model and the actual implementation model, making the outcomes of architectural analysis meaningless. To counter this, ArchJava [22] offers a unique binding between architectural description and actual implementation, but ArchJava does not yet provide indirect sharing verification.

Component contracts already were often proposed before for various purposes [23], [24]. For components written in Java, JML [25] is a popular formal contract specification language. The use of JML or related languages such as Spec# [31] for verifying component properties is a very active research domain. For example, Jacobs et al. [26] verify the absence of data races, and Pavlova et al. [27] focus on the security properties of applets. Other applications of JML are surveyed in [28].

The research presented in this paper proposes a pragmatic solution to broken session dependencies in Web applications. The main advantage of such pragmatism is the potential for short-term applicability. However, of course, research on more fundamental approaches is also needed and can have a more substantial impact in the long term. There is a large body of research on how to improve programming languages for programming distributed applications. In his keynote speech at ICSE 2005, Cardelli discussed three important areas where improvements are important: asynchronous concurrency, dealing with semistructured data and additional security abstractions [29]. The programming language E [32] is an example of a language that has emphasized security in its design. Other languages focus on specific classes of Internet applications such as distributed consensus applications [30].

IV. REQUIREMENTS

In particular, this paper eliminates certain types of runtime errors (such as a `NullPointerException` or a `ClassCastException`) by giving a formal guarantee that the no-broken-data-dependencies property is not violated in a given composition.

Reducing certain types of runtime errors by formally verifying that a given composition does not violate the desired composition property certainly improves the reliability of the software composition, but, in order to be really useful, the following interoperability and usability criteria are important as well:

4.1. Interoperability

It is important that the proposed solution is interoperable with the existing Web infrastructure and does not interfere with other Web security solutions.

4.2. Usability.

In order to encourage wide adoption by developers, we also identified two important usability characteristics for the solution:

4.2.1. Limited overhead.

In order to be generally applicable, the introduced overhead for the software developer and software composer must be minimal, both in terms of additional workload and verification time.

4.2.2. Applicability to real-life applications.

The applicability of the proposed solution may not be limited to toy examples, but the proposed solution must also be more generally applicable to larger real-life applications.

V. OVERVIEW OF THE SOLUTION

In this section we specify tree-based session dependency and use static and dynamic verification to guarantee that no-broken-data-dependencies property. Tree-based dependency proves that no-broken-data dependencies property in two steps.

5.1. Session control testing

We are able to read the D page only after reading the B and A pages. Without accessing B and A pages we can't access a page D. This illustrates the preventing broken data dependencies property and it is preventing the NullPointerException.

5.2. Session expiration testing

Generally session will expire with the fixed time. So we should access the pages only within the limited time, otherwise the session will expired. Here the page C is already accessed but the session of that page has been expired. So, we are unable to access the F page even if the C page has been accessed. This will provide the security.

Next to the implementation, the deployment information and the runtime Web traffic are also used as input for our verification process.

The verification process consists of three steps. First, the interactions with the shared session repository are explicitly specified in component contracts and static verification is used to verify that each component implementation obeys its contract specification. Second, the no-broken-data-dependencies property is verified in each possible execution path within a user's session. To verify this property statically, an upper bound is defined for the client-server interactions, namely, the intended client-server protocol. Next, the property is verified under the assumption that the client-server interactions are prefixes of the intended client-server protocol. Finally, runtime policy enforcement is used to guarantee that only the Web requests that are prefixes of the intended client-server protocol are processed by the Web application. By combining these three verification steps, our solution ensures the no-broken-data-dependencies property in a given application.

We will now discuss each of the three steps in more detail in the following sections.

Step1:- Server-Side Specification and Verification

In order to specify a component's interactions with the shared session repository, each Web component is extended with an appropriate component contract. The contract is expressed in a problem-specific contract language, which is easy to understand for application developers.

Step2:- Application-Specific Property Verification

The no-broken-data-dependencies property is verified by checking all of the possible execution paths in a user's session. To verify the property statically, an upper bound is defined for the client-server interactions, namely, the intended client-server protocol. This is an upper bound for the nondeterministic interactions between the client and the server and includes all valid client-server

interactions that may occur in the application under normal circumstances. The intended client-server protocol can be expressed in various ways, such as a regular expression, an EBNF notation, or a labeled state transition system.

Step3:- Runtime Protocol Enforcement

Finally, since the no-broken-data-dependencies property is verified under the assumption that all Web requests obey the intended client-server protocol; this assumption needs to be enforced at runtime. This can be done by loading the protocol specification into a supporting WAF or extending the application with an appropriate filter. As a result, only prefixes of the intended client-server protocol are allowed to be processed by the Web application.

VI. IMPLEMENTATION

In this part, we have to take one session control demo to prove the composition property. The composition property is as follows.

6.1. No-broken-data-dependencies

No client request causes a data item to be read from the server-side shared session repository before it has actually been written. For each shared data read interaction, the shared data item that already has been written to the shared session repository is of the type that is expected by the read operation.

By using session control demo we can prove the no-broken-data-dependencies property. To prove this property we have to use tree-based dependency in session control demo.

6.2. Tree-based-dependencies:

In this we can access session or page depend on the trees. The basic idea of this is as follows: All sibling nodes are accessed only when from root node to sibling pages on the tree are accessed.

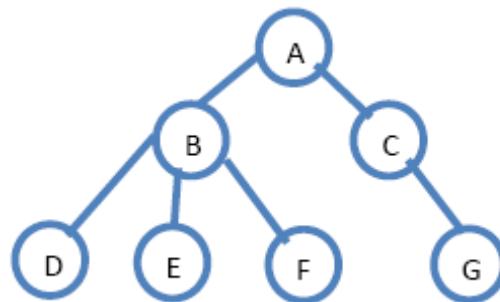


Fig.3. Tree-based dependency.

This tree-based dependency is used to prove the no-broken-data-dependency property. In this we have two steps.

First: we need to write session properties in the session control demo. We used these properties to write the tree-based algorithm. By using this algorithm we can see the process of accessing session pages. So first we need to write the session properties.

D	→	B, A
E	→	B, A
F	→	B, A
G	→	C, A
B	→	A
C	→	A

Second: next we have to write the session time out and data within the web.xml file because, based on this time the session will run. That means we can access the pages only within this time. Otherwise the session will expire. Then we can't access the pages.

```
<Session-config>
<Session-timeout>30000</Session-timeout>
..
</Session-config>
```

We are able to read the F page only after reading the B and A pages, and G page only after reading the C and A pages. This illustrates the preventing broken data dependencies property i.e. it is preventing the NullPointerException. Based on this session control demo I will implement the algorithm.

Algorithm Tree_based(x)

```

X ← node
S ← sibling node
P ← sibling node
R ← sibling node
If S ← X then
{
P ← S
R ← P
R ← accessed
P ← accessed
S ← accessed
}
else
R ← x
R ← accessed
```

This algorithm proves the tree-based dependency concept. In this we are testing the node or page in two cases. First, if the node can be sibling node, then we have to access the nodes from root to sibling. Second, if the node can be root node, we can directly access that node.

VII. RESULTS

We used tree-based verification experiment to measure the cost and verification performance of the proposed solution.

7.1. Tree-based verification performance

To evaluate the performance of the tree-based dependency verification process we used PPRWap framework. PPRWap framework is a provable protection for web applications PPRWap framework is a software framework, in computer programming is an abstraction in which common code providing generic functionality can be selectively overridden or specialized by user code providing specific functionality.

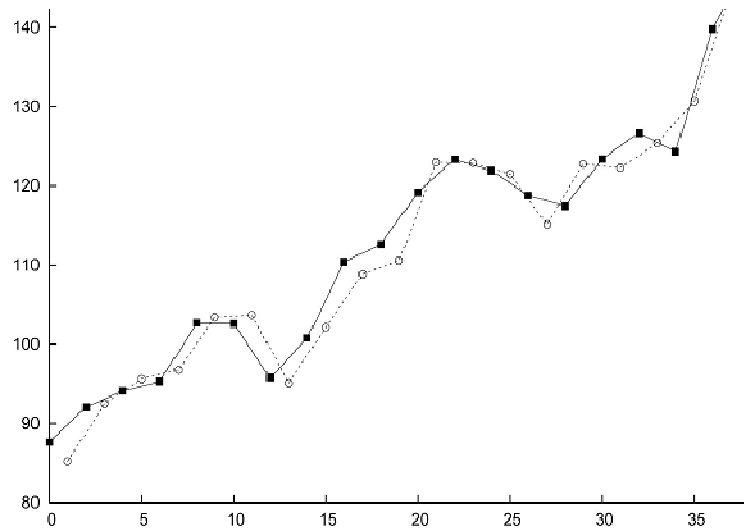

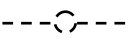


Fig. 4. Result of the overhead measurement.

Note:-
 Sequential dependency
 Tree based dependency

Generally tree-based techniques are easy to understand in any type of applications. By using tree and graphs we can easily understand session dependencies. So instead of sequential dependency, tree based dependency is better generally sequential dependency is very expensive and not easy to understand. So we used this tree-based dependency concept in session control demo with the help of the framework. The figure 2 shows the graphical representation between the sequential and tree-based dependency.

VIII. CONCLUSION

In this paper, we have presented an approach to prove that no-broken-data-dependencies property i.e. Tree-based dependency. This proof shows that the security and reliability of Web applications can be increased by guaranteeing in the absence of runtime errors. In particular, we have proposed a solution to prevent runtime errors due to broken data dependencies on session data. Our solution combines development-time program annotation, static verification, and runtime checking to provably protect against broken data dependencies in Web applications. We designed and developed a framework PP4Wap on session data. This will increase the security and reliability of data-centered Web applications.

Our solution also provides a good trade-off between usability and verification power. In this paper is to improve Web application security by providing an appropriate solution to the specific problem of broken data dependencies on session data.

REFERENCES

- [1] P.J. Deitel and H.M. Deitel "Internet and World Wide Web How to Program" Fourth Edition, Pearson International Edition, 2009.
- [2] T.E. Uribe and S. Cheung, "Automatic Analysis of Firewall and Network Intrusion Detection System Configurations," Proc. ACM Workshop Formal Methods in Security Eng., pp. 66-74, 2004.
- [3] Karl Forster, Lockstep Systems, Inc., "Why Firewalls Fail to Protect Web Sites," <http://www.lockstep.com/products/webagain/why-firewalls-fail.pdf>, 2007.
- [4] Sun Microsystems, Inc., "Java Servlet Technology," <http://java.sun.com/products/servlet/>, 2007.
- [5] V. Samar, "Unified Login with Pluggable Authentication Modules (PAM)," Proc. Third ACM Conf. Computer and Comm. Security, pp. 1-10, 1996.
- [6] E. Freeman, K. Arnold, and S. Hupfer, JavaSpaces Principles, Patterns, and Practice. Addison Wesley Longman, 1999.
- [7] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, "Hypertext Transfer Protocol—HTTP/1.1," IETF RFC 2616 (Category: Standards Track), <http://www.ietf.org/rfc/rfc2616.txt>, June 1999.

- [8] V. Raghvendra, "Session Tracking on the Web," *Internetworking*, vol. 3, no. 1, Mar. 2000.
- [9] M. Shaw and D. Garlan, *Software Architecture: Perspectives on an Emerging Discipline*. Prentice Hall, 1996.
- [10] I. Ristic, "Web Application Firewalls Primer," (IN) *SECURE*, vol. 1, no. 5, pp. 6-10, Jan. 2006.
- [11] webScurity, Inc., "The Weakest Link: Mitigating Web Application Vulnerabilities," http://www.webscurity.com/pdfs/webapp_vuln_wp.pdf, 2007.
- [12] V. Haldar, D. Chandra, and M. Franz, "Dynamic Taint Propagation for Java," *Proc. 21st Ann. Computer Security Applications Conf.* pp. 303-311, 2005.
- [13] A. Nguyen-Tuong, S. Guarnieri, D. Greene, J. Shirley, and D. Evans, "Automatically Hardening Web Applications Using Precise Tainting," *Proc. 20th IFIP Int'l Information Security Conf.*, R. Sasaki, S. Qing, E. Okamoto, and H. Yoshiura, eds., pp. 295-308, 2005.
- [14] C. Gould, Z. Su, and P. Devanbu, "Static Checking of Dynamically Generated Queries in Database Applications," *Proc. 26th Int'l Conf. Software Eng.*, pp. 645-654, 2004.
- [15] Y.-W. Huang, F. Yu, C. Hang, C.-H. Tsai, D.-T. Lee, and S.-Y. Kuo, "Securing Web Application Code by Static Analysis and Runtime Protection," *Proc. 13th Int'l Conf. World Wide Web*, pp. 40-52, 2004.
- [16] J. Offutt, Y. Wu, X. Du, and H. Huang, "Bypass Testing of Web Applications," *Proc. 15th Int'l Symp. Software Reliability Eng.*, pp. 187-197, 2004.
- [17] K. Golnabi, R.K. Min, L. Khan, and E. Al-Shaer, "Analysis of Firewall Policy Rules Using Data Mining Techniques," *Proc. 10th IEEE/IFIP Network Operations and Management Symp.*, pp. 305-315, Apr. 2006.
- [18] P.G. Neumann, "Keynote Speech: System and Network Trustworthiness in Perspective," *Proc. 13th ACM Conf. Computer and Comm. Security*, Oct.-Nov. 2006.
- [19] D.E. Perry and A.L. Wolf, "Foundations for the Study of Software Architecture," *ACM SIGSOFT Software Eng. Notes*, vol. 17, no. 4, pp. 40-52, 1992.
- [20] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*. Addison Wesley Longman, 1998.
- [21] N. Medvidovic and R.N. Taylor, "A Classification and Comparison Framework for Software Architecture Description Languages," *IEEE Trans. Software Eng.*, vol. 26, no. 1, pp. 70-93, Jan. 2000.
- [22] J. Aldrich, "Using Types to Enforce Architectural Structure," PhD dissertation, Univ. of Washington, Aug. 2003.
- [23] B. Meyer, "Applying 'Design by Contract'," *Computer*, vol. 25, no. 10, pp. 40-51, Oct. 1992.
- [24] B. Liskov, *Abstraction and Specification in Program Development*. MIT Press, 1986.
- [25] G.T. Leavens, "The Java Modeling Language (JML)," <http://www.jmlspecs.org/>, 2007.
- [26] B. Jacobs, K.R.M. Leino, F. Piessens, and W. Schulte, "Safe Concurrency for Aggregate Objects with Invariants," *Proc. Third IEEE Int'l Conf. Software Eng. and Formal Methods*, pp. 137-146, 2005.
- [27] M. Pavlova, G. Barthe, L. Burdy, M. Huisman, and J.-L. Lanet, "Enforcing High-Level Security Properties for Applets," *Proc. Sixth Smart Card Research and Advanced Application IFIP Conf.*, J.-J. Quisquater, P. Paradinas, Y. Deswarte, and A.A.E. Kalam, eds., pp. 1-16, 2004.
- [28] L. Burdy, Y. Cheon, D. Cok, M. Ernst, J. Kiniry, G.T. Leavens, K.R.M. Leino, and E. Poll, "An Overview of JML Tools and Applications," *Int'l J. Software Tools for Technology Transfer*, vol. 7, no. 3, pp. 212-232, June 2005.
- [29] L. Cardelli, "Transitions in Programming Models: 2," *Proc. 27th Int'l Conf. Software Eng.*, p. 2, 2005.
- [30] J.C.M. Baeten, H.M.A. van Beek, and S. Mauw, "Specifying Internet Applications with Dicons," *Proc. 16th ACM Symp. Applied Computing*, pp. 576-584, 2001.
- [31] M. Barnett, K.R.M. Leino, and W. Schulte, "The Spec# Programming System: An Overview," *Lecture Notes in Computer Science*, vol. 3362, pp. 49-69, Jan. 2005.
- [32] M.S. Miller, "Robust Composition: Towards a Unified Approach to Access Control and concurrency Control," PhD dissertation, Johns Hopkins Univ., May 2006.

Authors

Guda Shruthi received her B.Tech. in Computer Science and Engineering from Dr. VRK college of Engineering, Karimnagar, A.P, India and M.Tech. in Software Engineering from JITS, JNTU, Warangal, A.P, India. Currently, she is an Assistant Professor in the Department of Computer Science and Engineering, BITS, Warangal, A.P, India. She has more than 5 years of experience in teaching. Her research area of interest includes Computer networks, software engineering and neural networks.



Jayadev Gyani received his M.Tech. in Computer Science and Engineering from Osmania University, Hyderabad, India and Ph.D. in Computer Engineering from Central university of Hyderabad, India. Currently, he is a Professor in the Department of Computer Science and Engineering, JITS, Warangal, A.P, India. He has more than 16 years of experience in teaching. He also has Publications and presented more than 27 papers.



R. Lakshman Naik received his B.Tech. in Electronics and Telecommunication from SIET, JNTU, Krishna, A.P, India and M.Tech. in computer science and engineering from BITS, JNTU, Warangal, A.P, India. He also served as a System's Engineer in Wipro Technologies. He has Publications more than 7 papers in area of computer networks, Data mining and neural networks.



G. Sireesh Reddy received his M.Sc. in computers from KNR PG & Degree College Warangal, A.P, India and M.Tech. in computer science and engineering from BITS, JNTU, Warangal, A.P, India. He also served as a Software Engineer in Syntel India Ltd. His research area of interest includes computer networks.



ESTIMATION AND OPTIMIZATION OF PROSODIC TO IMPROVE THE QUALITY OF THE ARABIC SYNTHETIC SPEECH

Abdelkader CHABCHOUB & Adnen CHERIF

¹Signal Processing Laboratory, Science Faculty of Tunis 1060, Tunisia

ABSTRACT

The prosody modeling has been extensively applied in speech synthesis. This is simply because there is an obvious need for every speech synthesis system to generate prosodic properties of speech, for a natural and intelligible synthetic speech. This paper introduces a new technique for the prediction of a deterministic prosodic target at an early stage which relies on probabilistic models of F0 contour and may predict the duration. This paper, also, proposes a method that searches for the optimal unit sequence by maximizing a joint likelihood at both segmental and prosodic levels. This method has successfully been implemented in the analysis corpus for developing the Arabic prosody database which itself is the input of the Arabic speech synthesizer. This paper, also, shows a drastic improvement in the Arabic prosodic quality through extensive objective and subjective evaluation.

KEYWORDS: Segmental duration, pitch, predictive, prosodic Model, Neural Network, speech synthesis, Arabic speech.

I. INTRODUCTION

Generating natural sounding prosody is a central challenge in text-to-speech synthesis (TTS), which is nowadays a technology that enables computers to talk and assist people in learning languages. While existing synthesis techniques produce speech that is intelligible, few people would claim that listening to computer speech is naturally or expressive. Therefore in recent year's research in the areas of speech synthesis were directed more towards improving the intelligibility and natural of synthetic systems to achieve better quality, tone of voice as well as its synthetic speech and intonation [1] [2]. In several systems, the usability of systems-speech voice that produces a good quality still need extensive research to be able to increase its overall use.

In the Arabic language the processing linguistic and prosodic [3] is essential for the synthesis quality. So processing station based on the modification of the Arabic prosodic (optimization of the pitch and predictive duration) trained to improve the new Arabic voice. From the phonetic point of view, this is the processing of prosodic parameters defined by: the fundamental frequency (F0), segmental duration and intensity. Modeling of these parameters is the main target of our research work which essentially concentrates on the fundamental frequency and duration [4].

This paper is organized as follows. In Section 2, the morphological model of the Arabic language will be presented with, in particular the concepts of word. Section 3 describes the corpus used in the study and presents a list of phonemes and the corresponding acoustic parameters for each phoneme (duration and F0). These values are entered in the module to change the parameters that will optimize the prosodic parameters (pitch and duration) which will be presented in Section 4. Section 5 presents

the results and evaluation of the algorithm as well as the implementation of the speech synthesis system.

II. DATABASE OF ARABIC SPEECH PROSODY

The quality of a speech synthesis system depends on the intelligibility and naturalness of speech generated, hence the need for generating prosody quality. Our database has been developed to be used to improve the quality of Arabic synthetic speech with MBROLA [5].

The fundamental idea is to create a speech corpus consisting of phone sequence phonemic/prosodic context combinations that forms a specially structured subset of the set of all such combinations, and then use Arabic prosody transplantation [6].

The modules are cascaded in the order Phonetisation-Duration- Pitch. The input is a pair of a speech signal file and a time-aligned phonemic annotation, followed by phoneme validation (code SAMPA), followed by an identification of the voiced and un-voiced frames (V/NV), followed by duration extraction, followed by pitch extraction [7] and finally followed by Prosodic modification/optimization. This algorithm results are the entries of our Arabic prosodic database. The main data flow steps are shown in Figure 1 it represents the generation of the database.

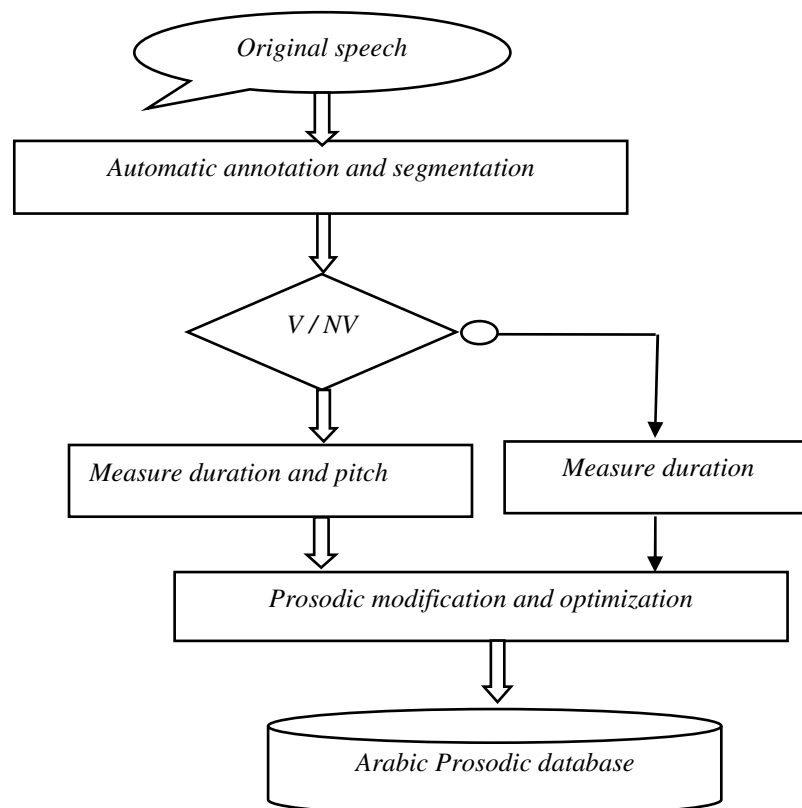


Figure1. Arabic prosodic database generation with prediction duration and pitch optimization algorithm.

2.1. Description of the corpus of analysis used

The corpus, which we used to build our database, is composed of 120 sentences, with an average of 5 words per sentence. These sentences contain in total 1296 syllables, 3240 phonemes, including a short vowels, long and semi- vowels [ا and ي], fricatives consonants, plosives and liquids consonants [ب and ط] and nasal consonants [م and ن]. Breaks were characterized with a "_" in the text corresponding to the natural voice. These sentences were read at an average speed (from 10 to 12 phonemes / second) by a speaker who did not receive any specific instruction to avoid any influence that could affect spontaneity. These corpus were recorded with a 16-Khz sampling rate and encoding 16 bits.

2.2. Segmentation and labeling of the corpus

Continuous speech corpus has been segmented and labeled by a semi-automatic procedure, which involves the following steps [12]:

- Step 1: phonetic spelling manual transcription of each sentence using the SAMPA transcription system.
- Step 2: Automatic segmentations by Praat.

2.3. Automatic segmentation of the corpus.

The extraction of pitch is an important step. For a period (ms) of phonemes, we extract the pitch in several positions that will be the parameters of the input file for the MBROLA, resulting in a pitch extraction algorithm robust and accurate, providing a good quality synthetic speech.

2.4. Identify the voiced and un-voiced frames.

The automatic segmentation of a speech signal is used in order to identify the voiced and un-voiced frames. This classification is based on the zero-crossing ratio and the energy value of each signal frame.

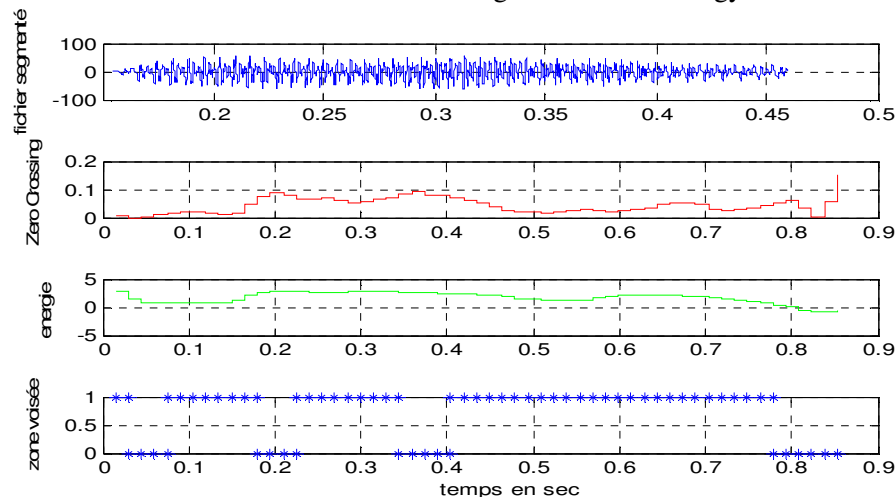


Figure2. Voiced zone Arabic Sentence "باب, door" This figure represents an automatic segmentation. For example, between [0.4, 0.75 (s)] as the voiced sections corresponds to a low Zero Crossing and high Energy.

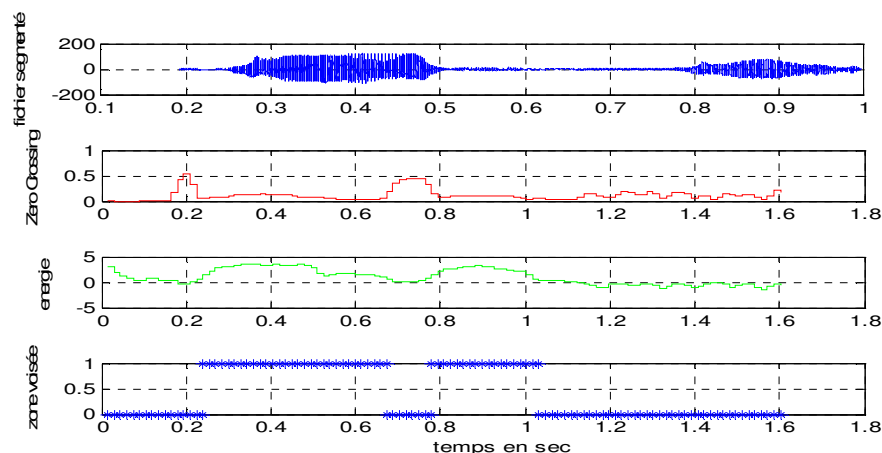


Figure3. Un-voiced zone, Arabic Sentence "شمس, sun" This figure represents an automatic segmentation. For example, between [0.65, 0.8 (s)] that the un-voiced sections corresponds to a high Zero Crossing and low Energy.

2.5. Duration and pitch extraction

The extraction of pitch is the next step. Copying the phonemes, the durations of the phonemes from the annotation file and measuring the pitch values from the original recording of a human utterance allows best case speech synthesis. To extract pitch from the recordings a Praat script called max-pitch was implemented as in [8], this script goes through the Sound and TextGrid files in a directory, opens each

pair of Sound and TextGrid, calculates the pitch maxima of each labeled interval, and saves the results to a text file [9]. The implementation of this script caused another problem and some modifications to the script were made. The inputs to this script are WAV files and TextGrid annotation files.

The Praat pitch extraction file produces one TXT file with the pitch values of all the phonemes in the files in the directory. The output “pitchresults.txt” file contains the following information:

1. File names of the files in the directory.
2. Labels.
3. Maximum pitch values of the labeled intervals in Hz.

The pitch results file for one file in a directory is shown in the following example: automatics extract Praat script .pho “صديقي العزيز”

```

_ 387
s. 90 83 132
a. 104 14 126 29 123 43 119 58 116 72 120 87 120
d 118 25 133 38 137 51 145 64 152 76 154 89 155
i 77 19 153 39 149 58 145 78 141 97 135
q 125
i 63 24 129 48 129 71 127 95 119
l 103 15 122 29 129 44 125 58 124 73 125 87 120
H 71 21 113 42 111 63 111 85 112
a 75 20 117 40 119 60 119 80 116
z 116 13 109 26 109 39 108 52 108 65 109 91 122
i 155 10 136 19 139 29 141 39 141 48 141 97 138
z 152 10 135 20 133 30 129 39 123 49 119 99 122

```

III. ARABIC PROSODIC MODELLING

3.1. Prediction models of segmental duration

Study analysis on the automatic generation of time have experienced many changes in recent years. The model proposed in this paper is based on two basic criteria that are: linear prediction and neural networks.

The model of W. N. Campbell assumes that the temporal organization of a statement is made at a higher level in terms of phonemes. Two stages are distinguished in the implementation of this model, the first is the prediction of syllabic duration and the second is the prediction of syllable phoneme durations.

A learning process automatically allows the prediction of syllable durations. It uses neural networks for learning because it is assumed that they can learn the fundamental interactions between contextual effects. These should represent the behavior governed by rules that are implicit in the data. If the networks can encode the fundamental interactions, then they would do the same with data not previously encountered [10].

Regarding the segmental durations, their distribution is given by the calculation of a coefficient of elongation (deviation from the mean). Campbell has suggested that all the phonemes of one syllable have the same elongation factor z : z -score. The z -score of each phonemic realization of the corpus of study is calculated by:

$$z_{realisatio\ n} = \frac{(dureeobser\ vee_{realisatio\ n} - \mu_{phoneme})}{\sigma_{phoneme}} \quad (1)$$

Where $\mu_{phoneme}$ and $\sigma_{phoneme}$ are the mean and standard deviation obtained from the absolute time of the achievements of each phoneme in the corpus. So, every time a phonetic realization is normalized by using the z -score (mean = 0 and standard deviation = 1) the durations of the syllables will be determined by the neural network[11]. Moreover, the model will calculate the z -score associated with each syllable by solving the following equation:

$$Duree(syllabe) = \sum_{i=1}^n \exp(\mu_i + z\sigma_i) \quad (2)$$

The sum on the phonemic elements of the syllable, z is the z -score associated with that syllable and the pair $(\mu_i$ and $\sigma_i)$ contains the mean and standard deviation associated with the phoneme i and obtained from the logarithms of the durations of achievements (in milliseconds) of this phoneme in the corpus. Thus, the duration of each phoneme of the syllable is calculated using equation (3).

$$Duree(phoneme_i) = \exp(\mu_i + z\sigma_i) \quad (3)$$

3.2 F₀ Prediction Module Based on a Neural Network

Neural networks provide a good solution for problems involving strong non-linearity between input and output parameters, and also when the quantitative mechanism of the mapping is not well understood. The use of neural networks in prosodic modeling has been reported in [13] and [14], but those methods do not make use of a model to limit the degrees of freedom of the problem. Additional care must be taken in order to account for the continuity of F₀ contours (using recurrent networks). In the proposed model, the continuity and basic shape of F₀ contours are ensured by the F₀ model [15][16].

In this paper, three types of neural network structures are evaluated: the multi-layer perceptron (MLP), Jordan (a structure having feedbacks from output elements), and Elman (a structure having feedbacks from hidden elements). The latter two neural network structures are called partial recurrent networks, and are tested here in order to account for the mutual influence of neighboring accentual phrases.

All structures have a single hidden layer containing either 10 or 20 elements. For the experiments, we utilized the SNNs neural network simulation software [17]. The results of F₀ contour prediction on the test data set are shown in Figure.4.

Figure.6 shows the pitch contour of an original and synthetic speech used with our system.

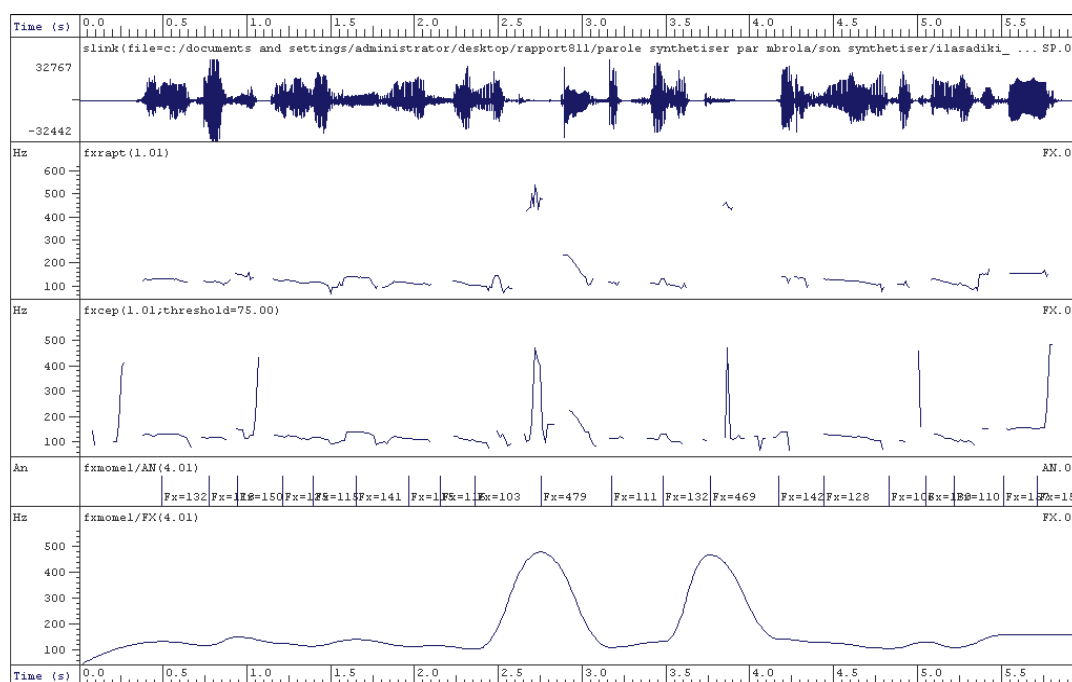


Figure 4. Evaluation of the fundamental frequency F₀ of the Arabic phrase "الى صديقي صديقي العزيز محمد كيف حالك" top-down voice signal, the varieties of F₀ autocorrelation method, spectral method, annotation segment average F₀ value of syllable and F₀ estimation by MOMEL.

IV. RESULTS AND EVALUATION

4.1 Implementation of prosodic values into MBROLA

MBROLA synthesis system is a multilingual; it was originally designed based on the characteristics phonostactics of the French language, our synthesis system requires for its adaptation to the Arabic language, with adjustments segmental and prosodic.

A first look at the results of the system showed that although there were similarities between the natural and synthetic versions, there is a considerable resemblance between the natural and synthetic F₀ contours. Only a few minor differences can be observed, since the F₀ values were extracted only once

every 10ms. Also note the halved F_0 in the creaky parts of the synthetic versions which successfully simulated creak. Similarly for the spectrogram there is small difference with the estimation algorithm. This can be seen in Figure.5 and Figure.6.

The implantation of our algorithm of estimation optimization of prosodic parameters produced an Arabic synthetic speech intelligible and natural.

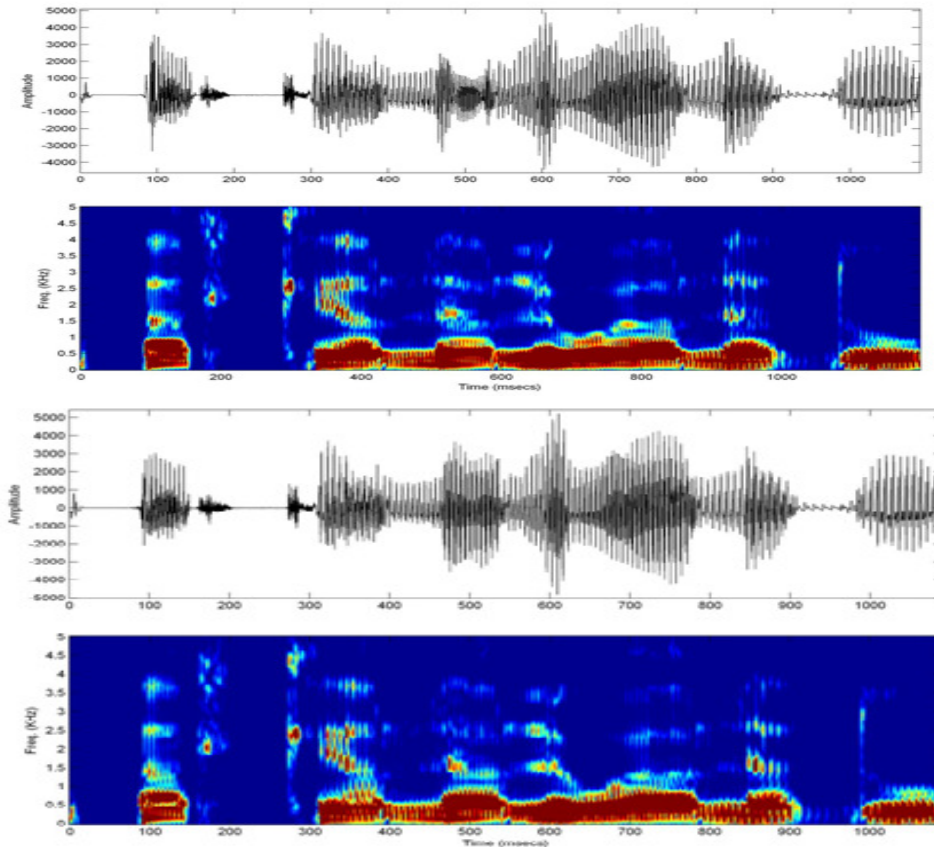


Figure5. Neutral and Synthesis speech, signal and spectrogram, Arabic Sentence «أكل الولد»

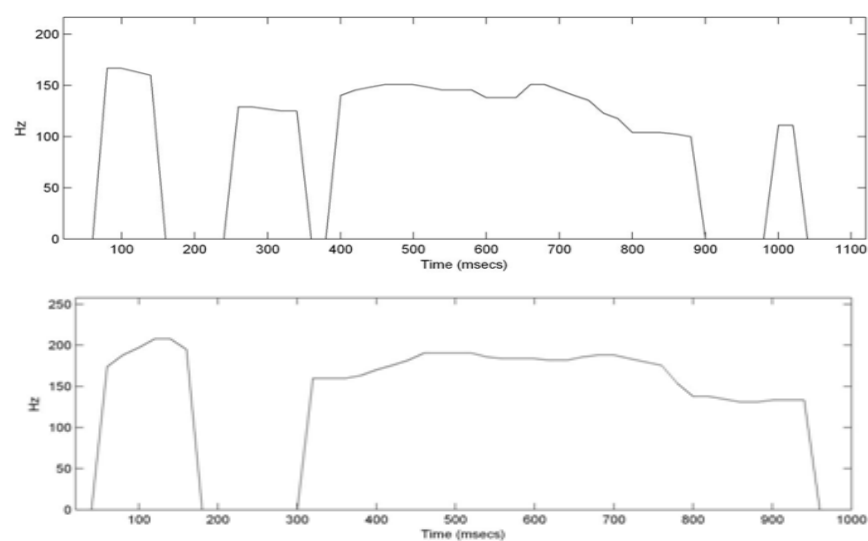


Figure6. Neutral and synthesis speech, pitch contour, Arabic Sentence, «أكل الولد».

4.2. Subjective evaluation

Evaluation consists of a subjective comparison between the 4 models. A comparison category rating (CCR) test was used to compare the quality of the synthetic speech generated by our system, Euler system, Acapela system and natural speech models. The listening tests were conducted by four Arab adults who are native speakers of the language. All listeners are born and raised in the Arab countries. For both listening tests we prepared listening test programs and a brief introduction was given before the listening test. They were asked to attribute a preference score according to the quality of each of the sample pairs on the comparison mean opinion score (CMOS) scale[18]. Listening test was performed with headphones. After collecting all listeners' response, we calculated the average values and we found the following results. In the first listening test, the average correct-rate for original and analysis-synthesis sounds were 98% and that of rule-based synthesized sounds was 90%. We found the synthesized words to be very intelligible Figure7.

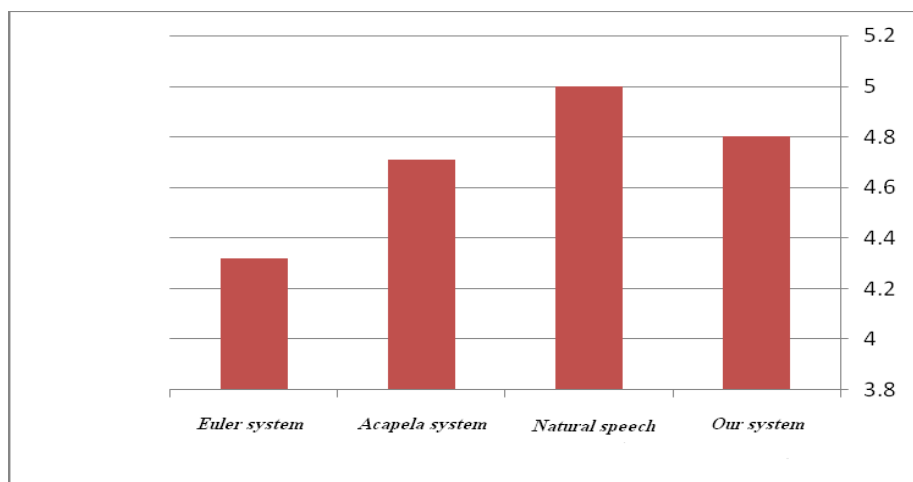


Figure7. Average scores for the first test (system Euler, our system, natural speech and Acapela system) for the intelligibility of speech.

V. CONCLUSIONS

A new high quality Arabic speech synthesis technique has been introduced in this paper. The technique is based on the estimation and optimization of the prosodic parameters such as pitch and duration for MBROLA method. It has also been shown in this paper that syllables produce reasonably natural quality speech and durational modeling is crucial for naturalness with a significant reduction in numbers of units of the total base developed. This was readily observed during the listening tests based on high quality and objective evaluation when comparing the original with the synthetic speech.

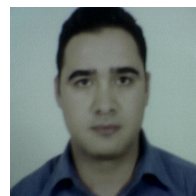
REFERENCES

- [1] S. Baloul, (2003) "Développement d'un système automatique de synthèse de la parole à partir du texte arabe standard voyellé", Thèse de doctorat, université du Maine, Le Mans, France.
- [2] M. Elshafi, H. Al-Muhtaseb M. Al-Ghamdi, (2002) "Techniques for high quality Arabic speech synthesis", Information Sciences 140-255-267, Elsevier.
- [3] M. Assaf, (2005) "A Prototype of an Arabic Diphone Speech Synthesizer in Festival", Master Thesis, Department of Linguistics and Philology, Uppsala University.
- [4] Möbius, B. and Dogil, G., (2002) "Phonemic and postural effects on the production of prosody", Speech Prosody 2002(Aix-en-Provence), p 523–526.
- [5] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, & O. van der Vrecken, (1996) The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers", Free of Use.
- [6] M. Al-Zabibi, (1990) "An Acoustic-Phonetic Approach in Automatic Arabic Speech Recognition", the British Library in Association with UML.
- [7] G. Demenko, S. Grocholewski, A. Wagner, & M. Szymański, (2006) "Prosody Annotation for Corpus Based Speech Synthesis", In: Proceedings of the Eleventh Australasian International Conference on Speech Science and Technology.

- Auckland, New Zealand, pp. 460-465.
- [8] Boersma, P. & Weenink, D. (2005) Praat. Doing phonetics by computer. [Computer program]. Version 4.3.04 Retrieved March 31, 2005 from <http://www.praat.org/>
 - [9] J. Bachan, & D. Gibbon, (2006) "Close Copy Speech Synthesis for Speech Perception Testing", In: *Investigationes Linguisticae*, vol. 13, pp. 9--24.
 - [10] W.N. Campbell, (1992) "syllable-based segmental duration", Edition G. Bailly and C. Benoît, *Talking Machines: theories, Models and Designs*, Elsevier Science Publishers, Amestrdam, pp.211-22 4,
 - [11] A. Lacheret-Dujour, B. Beaugendre, (1999) "La prosodie du français", Paris, Editions du CNRS.
 - [12] F. Chouireb, M. Guerti, M. Nail, and Y. Dimeh, (2007) "Development of a Prosodic Database for Standard Arabic", the *Arabian Journal for Science and Engineering*, Volume 32, Number 2B, pp. 251-262, ISSN: 1319-8025, October.
 - [13] S. Keagy. (2000) "Integrating voice and data networks: Practical solutions for the new world of packetized voice over data networks". Cisco Press,
 - [14] G. Sonntag, T. Portele and B. Heuft, (1997) "Prosody generation with a neural network: Weighing the importance of input parameters", in *Proceedings of ICASSP*, pp 931-934, Munich, Germany.
 - [15] J. P. Teixeira, D. Freitas and H. Fujisaki, (2003) "Prediction of Fujisaki model's phrase commands", in *Proceedings of Eurospeech*, Geneva, pp 397-400
 - [16] J. P. Teixeira, D. Freitas and H. Fujisaki, " Prediction of accent commands for the Fujisaki intonation model", in *Proceeding of Speech Prosody 2004*, Nara, Japan, March 23-26, 2004, pp 451-454.
 - [17] SNNS (Stuttgart Neural Network Simulator) User Manual (1995), Version 4.1, University of Stuttgart, Institute for Parallel and Distributed High Performance Systems (IPVR), Report No.
 - [18] K. S. Rao and B. Yegnanarayana, (4-8 October 2004) "Intonation modeling for Indian languages", in *Proceedings of Interspeech'04*, Jeju Island, Korea, pp733-736

Authors

A. Chabchoub: is a researcher in signal processing laboratory at the University of Sciences of Tunis – Tunisia (FST). Degree in electronics and he received a M.Sc. degree in Automatic and Signal Processing (ATS) from The National Engineering School of Tunis (ENIT). Currently, he is a PhD student under the supervision of Prof. A. Cherif. His research interests include speech synthesis and analysis.



A. Cherif: received his engineering diploma from the Engineering Faculty of Tunis and his Ph.D. in electrical engineering and electronics from The National Engineering School of Tunis (ENIT). Actually he is a professor at the Science Faculty of Tunis, Responsible for the Signal Processing Laboratory. He participated in several research and cooperation projects, and he is the author of international communications and publications.



INTEGRATION OF CONTROL CHARTS AND DATA MINING FOR PROCESS CONTROL AND QUALITY IMPROVEMENT

E.V.Ramana¹ and P. Ravinder Reddy²

¹Deptt. of Mech. Engg., Shadan College of Engineering & Technology, Hyderabad, India

²Deptt. of Mech. Engg., Chaitanya Bharati Institute of Technology, Hyderabad, India


ABSTRACT

Controlling the quality of materials, components, batches, products and assemblies during the course of manufacture is probably the most popularly recognized Quality Control activity. A charting and pattern recognition tool has been developed using Visual Basic and MS-Excel to automate the drawing of control charts. This can also enable better visualization of process data and provide automated recognition of basic unnatural patterns responsible for process to go out of control. Data mining has emerged as an important tool in extracting the knowledge from process databases. An out of control condition can trigger an appropriate data mining technique if any unnatural pattern is found in any control chart for a particular quality characteristic of any stage of the process. Process data can be analyzed by independent data mining technique to discover hidden patterns in the parameters that control the manufacturing process to improve the quality of products. Expert system database of the organization may be updated by the extracted knowledge from various stages of the process. An algorithm shall be developed that can control input parameters to various stages of the process so as to achieve its over all performance by making use of comprehensive expert system database of the organization along with the integration of control charts and data mining.

KEYWORDS: Data mining, control chart, Pattern recognition, data visualization

I. INTRODUCTION

Data mining is a collection of tools that explore data in order to discover previously unknown patterns. [16] Data mining applications include manufacturing process control, quality control, condition monitoring and fault diagnosis, optimization of manufacturing yield etc. It is a blend of techniques and concepts from statistics, AI, machine learning, neural networks, data visualization, pattern recognition etc.

Control chart may be defined as “a graphical method for evaluating whether a process is or is not in a state of statistical control”. It is a chronological (hour-by-hour, day-by-day) graphical comparison of actual product, part or other unit quality characteristics with limits reflecting the ability to produce as shown by experience on the unit-quality characteristics.[5] This paper focuses on application of control charts in monitoring the out of control condition that may occur at any stage of manufacturing process. If any out of control condition is found the reasons for the variation outside the stable patterns shall be discovered by appropriate data mining techniques like association rule, decision tree, regression, neural network etc and corrections are applied to input process parameters to control the product quality. This paper consists of six sections. In Section 2, definition of quality and causes for variation in quality are presented. In Section 3, types of control charts are described. Interpretation of control charts and basic tests that can be considered for pattern analysis on  chart are presented in Section 4. In Section 5, the methodology of integrating control charts and data mining for process control is presented. Finally the conclusions of this paper are presented in Section 6.

II. QUALITY

Quality is defined simply as meeting the requirements of the customer and this has been expressed in many ways

- Fitness for purpose or use
- The totality of features and characteristics of a product or product that bear on its ability to satisfy stated or implied needs

2.1 Variation in Quality

Every process and its output are subjected to variation. No two things can be made exactly alike. There are two types of causes responsible for change of dimensions (quality characteristics) on any component manufactured.

- Inherent or Random (chance) sources of variation: This is caused by large number of chance factors and depends on rigidity of machine tool and its process capability etc.
- Assignable causes of variation: This is caused by human factors, tool wear, machine wear, tool getting loose, difference in raw materials, environments which can be eliminated.

Quality control uses inspection as a valuable tool. In quality control programme inspection data are effectively used to take prompt corrective actions to arrest the incidence of defects in the job on hand and to plan the prevention of similar defects in future conceptually. The broad areas of application of quality control are incoming material control, process control and product control. [1, 13, 18]

III. TYPES OF CONTROL CHARTS

There are two fundamental types of control charts [5]

- Measurement or variables charts(of which the most popular are \bar{X} , R, S charts) for use when actual readings are taken
- Charts for use with go and no go or attributes are called p, np, c and u charts

Once the quality characteristic to be studied has been determined; data is collected and samples must be properly selected by the principle of rational sampling. Rational samples are groups of measurements, the variation among which is attributable to one system of causes. Examples of non rational sampling include sampling from parallel machines, sampling over extended period of time and sampling from products of several sources.

If the quality related problems are not easily solved in numerical form, then Attribute Control Charts (ACC) are useful. ACC tools are effective for detecting quality improvement or degradation. ACC is used to evaluate the variation in a process where measurement is an attribute i.e. is in discrete or count (e.g. pass/fail, number of defects). In ACC defining sample size is a problem. A fuzzy based model for ACC in multistage processes was developed and the model is solved by GAs (Genetic Algorithms) and the sample size is suggested. [19]

IV. INTERPRETATION OF CHARTS

- When interpreting charts, it is important to start with the R chart and get it under statistical control first because the limits of \bar{X} chart depend on the magnitude of the common cause variation of the process measured by \bar{R} . If some points on the R chart are initially out of control, the limits on \bar{X} chart will be inflated.
- The occurrence of trend or recurring cycles in the data pattern can indicate that system experiencing a drift or cyclical change with respect to its mean or range
- Run of points above or below the center line may be indicating small shifts in the mean or level of variability

When ever an out of control condition is indicated, it is important to determine the basic process fault that is producing it. Some useful generic conditions to look for include the following:

- Trends/cycles :- Systematic change in the process environment, worker fatigue, maintenance schedules, wear conditions, accumulation of waste materials and contamination

- Sudden shifts in level :- New machine, die or tooling, new batch of material, change in measurement system and change in production method

The unusual patterns were fairly obvious even through the most careful examination of the charts. Often, special causes of variation produce patterns that are less obvious. Therefore more rigorous pattern examination should generally be conducted. Several useful tests may be performed by dividing the distance between upper and lower control limits into six zones as shown in Fig-1. Each zone is one standard deviation wide. The zones are termed as Zone A, Zone B, Zone C from the upper control limits toward the central line. The lower side is reflected as mirror images. The probabilistic basis for the tests to be discussed using the zones is derived from the normal distribution. Basic tests that can be considered for pattern analysis on \bar{X} chart are presented in Fig.2. However some tests are applicable for attribute charts that don't follow the normal distribution. Test 1 works for all types of attribute charts. More over Test 2, Test 3 and Test 4 are applicable for p chart. [8]

LEGEND

UCL: Upper Control Limit

UWL: Upper Warning Limit

CL: Center Line

LCL: Lower Control Limit

LWL: Lower Warning limit

+SL: Upper Sigma Limit

-SL: Lower Sigma Limit

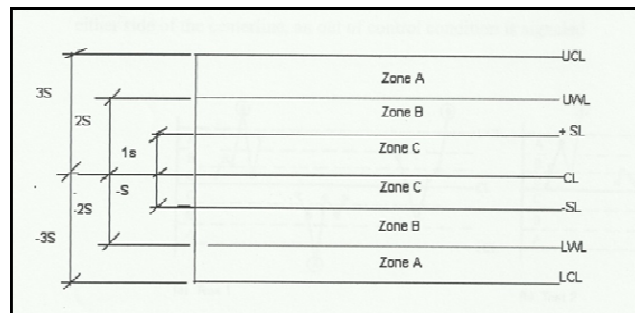


Fig-1 Zone Rules for control chart analysis

Test 1: The existence of a single point beyond Zone A signals the presence of an out of control condition. (Extreme points)

Test 2: Two out of three points in Zone A or beyond signals the presence an out of control condition

Test 3: Four out of five points in zone B or beyond signals the presence of an out of control condition

Test 4: Seven or more successive points either strictly above or strictly below the center line (this rule is applicable to both \bar{X} and R charts)

Test 5: When six successive points on either \bar{X} or R charts show a continuing increasing or decreasing trend, a systematic trend in the process is signaled

Test 6: When 14 successive points, oscillate up and down on either \bar{X} or R charts, a systematic trend in the process is signaled

Test 7: When eight successive points, occurring on either side of the center line with none in Zone C, an out of control condition is signaled

Test 8: When fifteen successive points on the \bar{X} chart fall in Zone C only, to either side of the center line, an out of control condition is signaled

Although these tests can be considered as basic, they are not totally comprehensive. Analyst should be alert to any patterns of points that might indicate the influence of other special causes in the particular process. Tests 1, 2, 5 and 6 are separately applied to the upper and lower halves of the chart. Tests 3, 4, 7 and 8 are applied to the entire chart (\bar{X} and R charts). Points can contribute to more than one test. [1, 8, 13, 18]

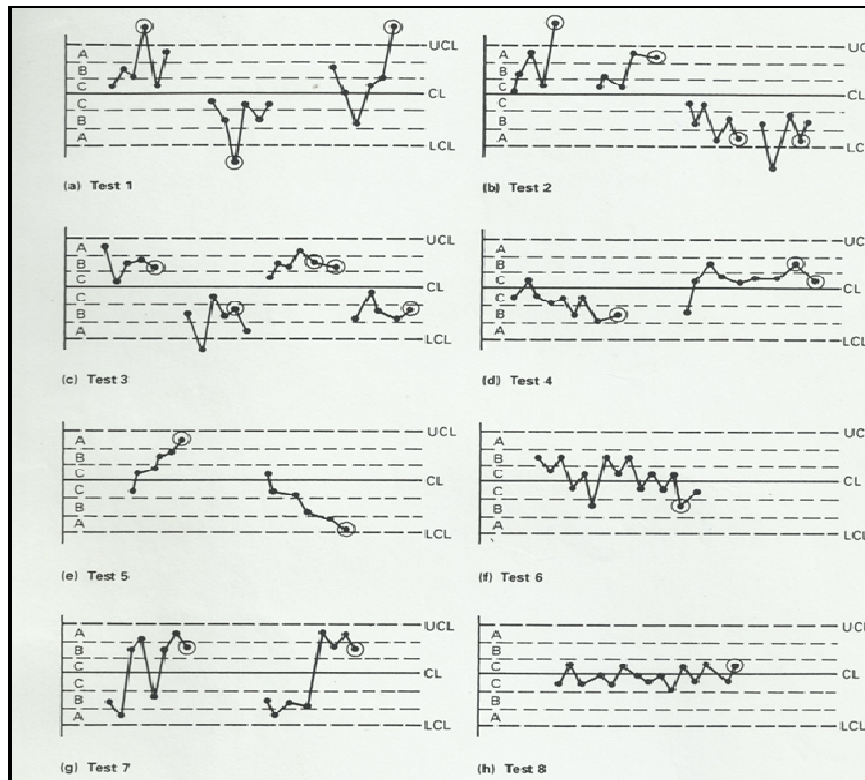


Fig-2 Pattern analysis of \bar{X} charts

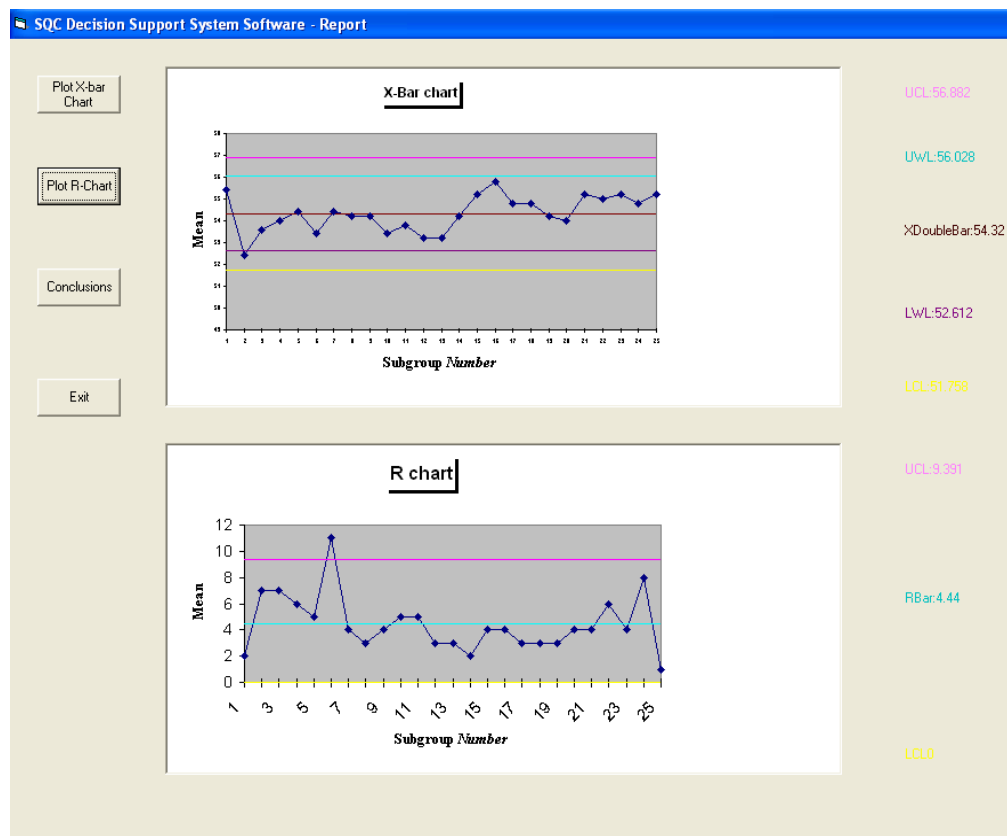
V. INTEGRATION OF CONTROL CHARTS AND DATA MINING FOR PROCESS CONTROL

To plot control charts and drawing conclusions from them requires collection and organizing the process data, determining control limits, recognition and analysis of unusual patterns exhibited by the process. More over plotting the charts manually for different quality characteristics becomes tedious work and error prone. This consumes more time and requires greater attention. A charting and pattern recognition tool has been developed using Visual Basic and MS-Excel [7, 9] to automate the drawing of control charts. This tool provide better visualization of process data and enable automated recognition of basic unnatural patterns responsible for process to go out of control.

Typical data entry sheet for manual collection of sample data to draw \bar{X} and R charts is shown in Fig-3. \bar{X} and R charts generated by the charting tool is presented in Fig-4 for typical data. (XDoubleBar: Arithmetic mean of the sample averages, RBar: Arithmetic average of sample ranges) Pattern recognition tool performs tests on collected data to discover basic unusual patterns responsible for unstable process. Passing of a test indicates absence of unusual pattern linked with that test and vice versa. A typical output showing the outcome of tests performed on \bar{X} and R charts is given in Fig-5.

SQC Decision support system software - XBarChart-Data Entry Sheet

	A	B	C	D	E	F
1	Next>	Measured Value1	Measured Value2	Measured Value3	Measured Value4	Measured Value5
2	SubGroup1	54	56	56	56	55
3	SubGroup2	51	52	54	56	49
4	SubGroup3	54	52	50	57	55
5	SubGroup4	56	55	56	53	50
6	SubGroup5	53	54	57	56	52
7	SubGroup6	53	47	58	55	54
8	SubGroup7	52	55	54	55	56
9	SubGroup8	56	53	53	54	55
10	SubGroup9	55	52	53	56	55
11	SubGroup10	50	54	53	55	55
12	SubGroup11	57	54	53	52	53
13	SubGroup12	52	52	54	53	55
14	SubGroup13	54	53	55	52	52
15	SubGroup14	54	55	54	53	55
16	SubGroup15	56	53	57	56	54
17	SubGroup16	58	57	56	54	54
18	SubGroup17	55	55	55	56	53
19	SubGroup18	54	57	54	55	54
20	SubGroup19	54	53	56	53	55
21	SubGroup20	53	53	57	54	53
22	SubGroup21	53	55	57	56	55
23	SubGroup22	59	54	53	54	55
24	SubGroup23	54	55	58	55	54
25	SubGroup24	56	53	51	55	59
26	SubGroup25	56	55	55	55	55

Fig-3 Data entry sheet for \bar{X} and R chartsFig-4 \bar{X} and R charts generated by the charting tool

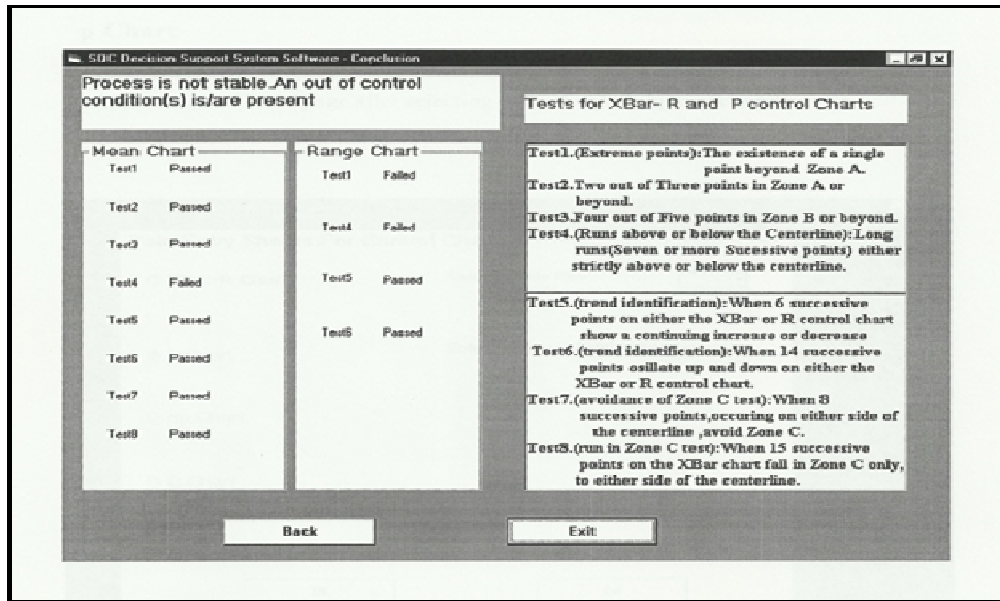


Fig-5 Outcome of tests performed on \bar{X} and R charts

Quality troubles often gradually “drift” into a process. Tool wear deflection, tool temperature or an improperly ground tool may result in a trend which finally leads to production of non-conforming parts. A chart which compares actual production variations with control limits may flag the entrance into the process of this sort of quality trouble before scrap or rework is actually caused. Unusual variation caused by assignable causes may represent temporary difficulties that can be eliminated with out much expenditure. [5]

Classical methods like control charts aim to monitor the process not to infer the relationship between target attribute and input attributes. [16].The goal of mining quality related data is to find the relation between the quality measure (target attribute) and input attributes (manufacturing process data). The idea is to predict the quality measure of a certain product/ batch based on its process parameters.

The control signature is a set of feature values or ranges that lead toward an expected output. Rough set theory was used by Kusiak to determine the association between control parameters and product quality in the form of decision rules and generated control signature from those rules.[15] The manufacturing parameters include the characteristics of the production line such as machine that has been used in each step, how machine has been set up, raw material in the process, environment (moisture, temperature), operators on the production line (experience level of worker, which have been assigned on each machine in the line), the shift number and other significant factors. [16]

Data is collected electronically and stored in databases in modern plants. The volume of collected data makes it impractical to explore and detect intricate relations between parameters of different processing steps using standard statistical procedures. Also comprehensive manual analysis of voluminous data becomes prohibitive, especially on-line monitoring and control of manufacturing processes are considered.

Data can be analyzed by data mining techniques to identify hidden patterns in the parameters that control manufacturing processes or to determine and improve quality of products. [2] “Interestingness” (of pattern) is usually taken as an overall measure of pattern value, combining validity and novelty, usefulness, simplicity and understandability. The measures of interestingness can be classified into objective measure and subjective measure. An objective measure depends on structure of pattern and underlying data used. Subjective measure depends on the class of users who examine the patterns. These are based on two concepts. [17] Unexpectedness (a pattern is interesting if it is unexpected) and actionability (a pattern is interesting if the user can use it to his or her advantage) of the pattern.

Study of entire process data/factory data altogether to discover the problem areas instantly affecting subsequent processes requires consideration of hundreds of attributes simultaneously. During the

product design and development process data mining can be used in order to determine “internal” factors at each stage and “external” factors at consecutive and previous stages. Data mining can be seen as a supporting vehicle for determining causal relationships among “internal” factors associated with manufacturing processes and external factors related to the competitiveness of the manufacturing company. Since manufacturing have an inherently temporal context, time factors may be taken into process in order to correctly interpret the collected data. (e.g., from a certain production date the number of defects is much higher than normal)

The deviation of process data from the normal and/or discernible pattern detected over time may have an adverse impact on current and subsequent stages. Extracted knowledge may include classification to predetermined types of deviation from the norm and causal relationships among temporally oriented events. [6]

Different data mining techniques may be applied with independent goals at each stage of the process. The failed batch / yield identified by a date can help in identifying the cause relating to faulty equipment or tool or material etc. Condition based monitoring information could come from either or both on-line condition monitoring data and historical records. Process analysis is concerned with optimizing inter related processes with in an enterprise. Identifying inter relationships among these processes has proven a valuable field of data mining technology. The knowledge extracted by mining process databases at various stages can be integrated with existing expert systems of the organization. Various decision making tools can be networked in support of globally optimal decisions. [3] Data mining tools have shown a profound impact on informed, transparent and autonomous decision making. An algorithm can be developed to integrate overall system performance with the component systems based on extracted knowledge.

Control chart helps in monitoring a particular stage of a process and necessary for the development of process know how. It can trigger a specific data mining technique applicable for that stage if any out of control pattern is detected and expert system database of the organization can be updated with the extracted knowledge. Fig-6 shows the integration of control charts and data mining for process control.

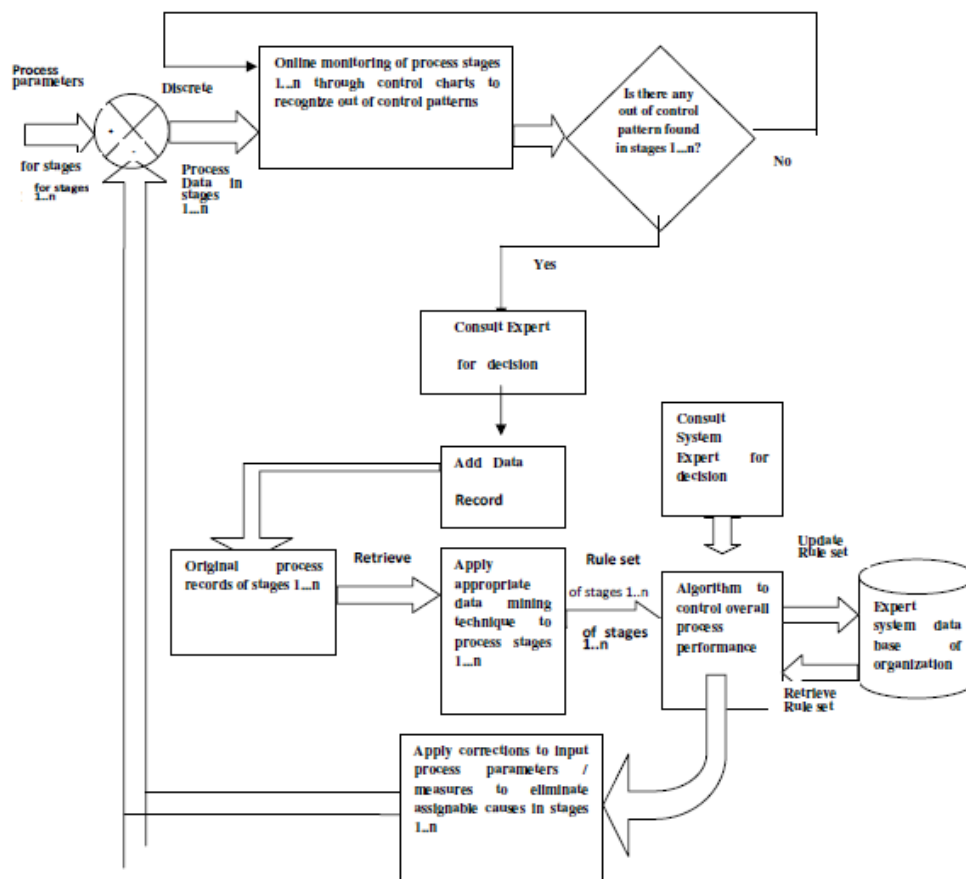


Fig-6 Integration of control charts and data mining for process control

VI. CONCLUSIONS

This paper presented the methodology for the integration of control charts with data mining techniques to achieve process control. Knowledge may be extracted by applying data mining techniques on temporal data collected at various stages of the process; condition based monitoring data of equipment, tools etc and analysis of defects in products/components produced.

Expert system database of the organization may be updated by the extracted knowledge from various sources as mentioned above. An algorithm can be developed to integrate over all process performance (e.g. performance parameters, production indices, yield, company goals etc) with its component stages through comprehensive expert system database. The algorithm is expected to provide optimal process parameters (input) to each stage of the process so as to improve over all process performance.

References

- [1] "ASM Handbook: Non destructive evaluation and quality control", ASM International, pp728-732, 1989
- [2] A.K.Choudhary, J.A.Harding, M.K.Tiwari, Data mining in manufacturing: a review based on kind of knowledge, Journal of Intelligent Manufacturing, Vol. No: 20, pp501-521, 2008
- [3] A.Kusiak, "Data mining: manufacturing and service applications", International Journal of Production Research, Vol. No: 44, pp4175-4191, 2006
- [4] Alex G.Buchner, Sarabjot S.Anand and John G.Hughes, "Data Mining in Manufacturing Environments: Goals, Techniques and Applications", University of Ulster, Northern Ireland, UK
- [5] Armand V.Feigenbaum, "Total Quality Control", McGraw-Hill International Editions, pp 395-409, 1991
- [6] Dan Braha (Ed), Data mining for Design and Manufacturing: methods and applications, Kluwer Academic Publishers, 2002, ISBN: 1-4020-0034-0
- [7] Evangelos Petroutsos, "Mastering Visual Basic 6.0", BPB Publications, pp10-65, 2001
- [8] Gavriel Salvendy, "Hand book of Industrial Engineering", John Wiley & Sons, pp2259, 1992
- [9] Gene Weisskopf, "Excel 2000 No Experience Required", BPB Publications, pp54-81, 2003
- [10] Hovhannes Sadoyan, Armen Zakarian, Pravansu Mohanty, "Data mining algorithm for manufacturing process control", The International Journal of Advanced Manufacturing Technology, Vol. No: 28, pp342-350, 2006
- [11] J.A. Harding, M.Shahbaz, Srinivas, A.Kusiak, "Data mining in Manufacturing: A Review," Journal of Manufacturing Science and Engineering, Vol. 128, pp968-976, 2006
- [12] Jing Rong Li, Li Pheng Khoo, Shu Bengtor, "RMINE: A rough set based data mining prototype for the reasoning of incomplete data in condition based fault diagnosis", Journal of Intelligent Manufacturing, Vol. No: 17, pp 163-176, 2006
- [13] John S.Oakland, "Statistical Process Control", Butterworth-Heinemann Ltd, 1999
- [14] Kescheng Wang, "Applying data mining to manufacturing: the nature and applications", Journal of Intelligent Manufacturing, Vol. No: 18, pp487-495, 2007
- [15] Kusiak, A., "Data mining approach for generation of control signatures" ASME Journal of Manufacturing science and Engineering, Vol. No: 124, pp 923-926, 2002
- [16] Lior Rokach, Oded Maimon, "Data mining for improving the quality of manufacturing: a feature set decomposition approach", Journal of Intelligent Manufacturing, Vol. No: 17, pp285-299, 2006
- [17] Silberschatz, A., Tuzlililn, A., "What makes patterns interesting in knowledge discovery systems", Knowledge and Data Engineering, vol. No: 8, pp970-974, 1996
- [18] Tom Drozda, Charles Wick, John T. Benedict, Raymond F.Villeux, "Tool and Manufacturing Engineers Handbook("Quality control and assembly)", Society of Manufacturing Engineers, pp2-17, 1986
- [19] Orhan Engin, Ahmet C, elik, I'hsan Kaya, "A fuzzy approach to define sample size for attributes control chart in multistage processes: An application in engine valve manufacturing process", Applied Soft Computing, Vol:8, 2008, pp1654-1663

Authors

E. V. Ramana is working as a Professor & Head in the Department of Mechanical Engineering, Shadan College of Engineering & Technology, Hyderabad. He received his first M.Tech degree in Energy Systems and second M.Tech degree in CAD/CAM in the years 1992 and 1997 from JNT University, Hyderabad. He possesses 14 years of experience in teaching and 6 years in industry and research. He is currently pursuing the PhD degree from JNTUH, Hyderabad, India. He has considerable experience in developing mechanical engineering related application soft wares. His active area of research is Data Mining in Manufacturing.



P. RAVINDER REDDY is born in 1965 and he is working as a Professor and Head of Mechanical Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad. He is having 22 Years of Teaching, Industrial and Research experience. Taught Postgraduate and under graduate Engineering subjects. Published Research Papers over 132 in International and national Journals, and Conferences. Guided 5 Ph.Ds and 6 Ph.D scholars submitted their thesis. Guided over 250 M.E/M.Tech Projects and carried out research and consultancy to a tune of Rs. 1.9 Cr sponsored by BHEL, AICTE, UGC, NSTL and other industries. Organized 23 Refresher/STTPs/ workshops, one international conference and delivered 63 invited/ keynote/ special lecturers. Received “UGC Fellowship” award by UGC (1999). Raja Rambapu Patil National award for promising Engineering Teacher by ISTE for the year 2000 in recognition of his outstanding contribution in the area of Engineering and Technology. Excellence “A” Grade awarded by AICTE monitoring committee for the MODROB project sponsored by AICTE in 2002. “Engineer of the year Award-2004” for his outstanding contribution in Academics and research by the Govt. of Andhra Pradesh and Institution of Engineers (India), AP State Centre on 15th September 2004 on the occasion of 37th Engineer’s Day. Best Technical Paper Award in the year Dec. 2008 by National Governing Council of Indian Society for Non Destructive Testing.



BINS APPROACH TO IMAGE RETRIEVAL USING STATISTICAL PARAMETERS BASED ON HISTOGRAM PARTITIONING OF R, G, B PLANES

H. B. Kekre¹ and Kavita Sonawane²

¹Professor, Department of Computer Engineering NMIMS University,
Mumbai, Vileparle, India

²Ph.D Research Scholar NMIMS University,
Mumbai, Vileparle, India

ABSTRACT

In this paper we have proposed a novel technique to retrieve the images from large image databases based on the spatial contents of the image to extract the low level features from it for CBIR. In this work image is separated into 3 planes and for each plane we have calculated the histogram which is partitioned into three equal parts to obtain the 27 bins. These bins are holding the spatial-color information of the image in various forms which is generating the different types of the feature vector databases. Three different set of bins are designed to be used as feature vectors containing the total R,G, and B intensities, in R, G and B bins respectively, second form is containing the mean of R,G and B and third is holding standard deviation of R, G and B values in R, G, B bins respectively. This leads to generation of three feature vector databases where size of each feature vector in all databases is 27. Experimentation includes comparison of 100 query images with 1000 database images (Augmented Wang database) using two similarity measures named Euclidean distance and Absolute distance. Results obtained for three feature databases are compared based on the similarity measures reflecting the performance variations of different approaches used. These results are then analyzed and refined using three criteria in this work named Criterion1: Strong, Criterion2: Average and Criterion3: Weak. It has been observed in our results that bins holding standard deviation of R, G and B intensities performing better among all three approaches and refined results using criterion3 giving very good results as compared to other two.

KEYWORDS: CBIR, Histogram Partitioning, Mean, Standard Deviation, Euclidean Distance, Absolute Distance, LIRS, LSRR, PRCP.

I. INTRODUCTION

This paper explores the new and effective technique to retrieve the images from large image databases using the image contents. Early image retrieval methods locate the desired images by matching keywords that are assigned to each image manually. However because of large number of images it is impractical to assign keywords to every image and even difficult to process them manually. Now a days, content based image retrieval seems to be very popular and fastest growing research area. In all the CBIR system the overall work can broadly be divided into two parts, First, Feature extraction and Second, selecting and applying the similarity measure which will retrieve the image closer to query with less complexity and good accuracy in terms of precision and recall [1],[2]. Most of the CBIR systems are using the primitive features of the images denoting the visual information including the color, shape and texture features [3],[4],[5],[6]. Color is independent of image size and orientation, as it is robust to background complication.[7],[8],[9],[10],[11]. We are focusing on the use of color information of the image obtained by calculating the image histograms to extract the feature vectors [12].

Histogram is most commonly used technique to represent global feature of the image which is invariant to translation and rotation of the image [13],[14],[15],[16]. Histogram for an image is constructed by counting the number of pixels of each color intensity level. [17], [18], [19],[20],[21]. If we consider the histogram as a feature vector directly for comparison process, although it is small vector representing the image it would be better if further we can reduce the size of the feature vector so that comparison computations can be minimized and complexity can be reduced [22],[23]. Keeping this goal in mind we have calculated the histogram of each plane of the image which is divided into three equal parts. This partitioning of the histogram is based on dividing the number of pixels equally in three parts. The intensity where this partition takes place is taken into consideration on the basis of which actual 27 bins are formed to be considered as feature vectors. Here three feature vector databases are formed based on three different parameters. Feature extraction from each of these bins is nothing but the color information representation in three different ways as the total intensities, average intensities and the standard deviation of these intensities in each bin. As second part of CBIR, this system is tested using the query by image, given as input which is compared with the 1000 database image feature vectors using two similarity measures named Euclidean distance and Absolute distance [24],[25]. An image similar to the query where the distance is less is selected for retrieval. To evaluate the performance of the system instead of using the traditional parameters precision and recall as it is, we are using 'Precision- Recall Cross over Point (PRCP)' to evaluate our system's performance [26], [27]. Further these results are analyzed and refined using three simple criteria designed [28], [29]. Criterion 1, and 2 and 3 gives the gradual increase in the performance level of our system. Orientation of this proposed work is as follows: Section2 explains the feature extraction process, Section3 describes application of the similarity measures, experimental results and analysis, and performance evaluation of the system is given in section 4, Work done is concluded section 5.

II. FEATURE EXTRACTION

Feature extraction process starts with separation of image into R, G, B planes. Then computing the histogram of each plane which is divided into parts to generate the bin addresses and finally it ends with formation of bins to represent the feature vectors is explained below.

2.1. Histogram Plotting and Partitioning

As mentioned above each image for which the feature to be extracted is separated into R, G and B planes [32]. Histogram for each plane is computed separately. Once the histograms are obtained each is divided into 3 equal parts based on the total number of pixels in the image. This gives two separation points of three parts, where the intensity levels are taken into consideration as Grey Level 1 and 2 as GL1 and GL2 respectively. Each part of histogram is assigned an id as 0, 1 and 2. As shown in figure.1. Same process is repeated for 3 planes.

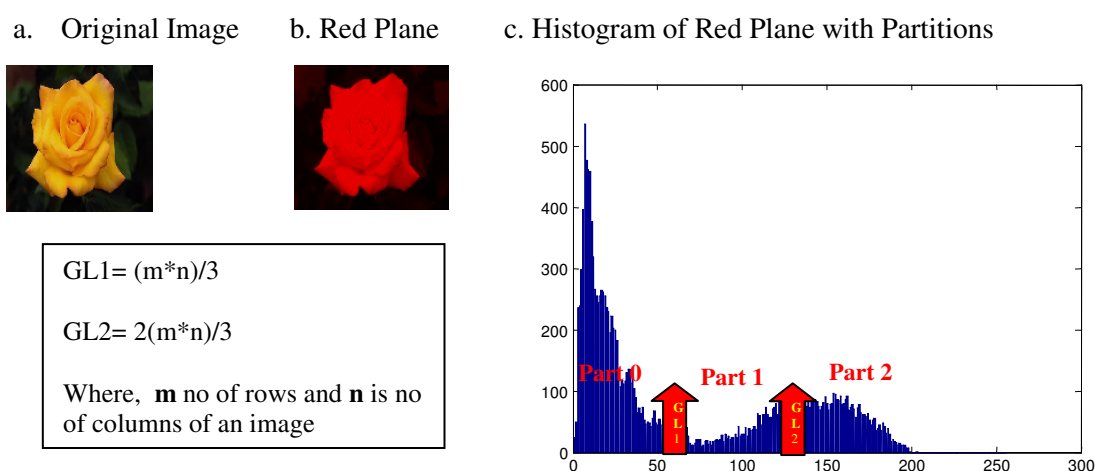


Figure1. a. Original Image b. Red plane c. Histogram with partitions (Red Plane)

2.2. Bins Formation

This process uses the information obtained from the histogram. As shown in Figure 1, the histogram gives two grey levels named as GL1 and GL2 for 3 partitions. These grey levels are used as threshold for the pixel to be counted in a bin which decides that in which part of the histogram it falls either 0, 1 or 2. According to this logic, for each pixel from the original image with its R, G, B values it will be checked that which part it belongs in the respective histogram and a flag will be assigned to it. This way the flag will have 3 values in range (000 to 222) total 27 possibilities as each pixel has three components R, G and B. This 3 digit flag assigned to each pixel actually generates the address of the bin for that pixel to reside or to be counted. The bin address can be identified using the following Equation1.

$$\text{Bin_Address} = (9r + 3g + b) \quad (1)$$

Where ‘r,’ ‘g’ and ‘b’ are the flag digits for R, G and B values respectively for the pixel in process

e.g if the flag values are r=2, g=1 and b=1 then that pixel will be counted in Bin_22.

2.3. Feature Vector Databases

Once the bins are formed as explained in 2.1 and 2.2, we thought of the roles these bins can play in terms of containing the information. Based on this we commanded the bins to carry pixel information in 3 different ways which classifies the bins into 3 formats. First is ‘Bin_Total_R_G_B’ holds the total red, green and blue intensities of the pixels counted in that particular bin separately. Second is ‘Bin_Mean_R_G_B’ carrying the mean R, G, and B values of the number of pixels in each bin and third format is ‘Bin_Standard_deviation_R_G_B’ which counts the standard deviation of the R, G, B intensities of the number of pixel in each bin separately. The way these parameter are computed is given equation 2, 3 and 4. As explained earlier total bins obtained are 27 from 000 to 222. and as we are considering R,G B separately along with the three parameters to represent the feature vector; we have multiple feature vector databases 3(R, G and B) for each of the three formats- Total, Mean and Standard deviation where each feature vector is having 27 components. Same process is applied to all 1000 images in the database and feature databases are prepared as pre-processing part of this work. Means in all there are 9 feature vector databases for which the comparison should take place.

Let there be N pixels in a bin. Each pixel having intensity level as Ri, Gi and Bi respectively for R,G and B planes.

$$\text{Bin_Total_R} \rightarrow R_T = \sum_{i=1}^N R_i \quad (2)$$

$$\text{Bin_Mean_R} \rightarrow \bar{R} = \frac{1}{N} \sum_{i=1}^N R_i = \frac{R_T}{N} \quad (3)$$

$$\text{Bin_Standard_deviation} \rightarrow R_{SD} = \frac{1}{N} \sqrt{\sum_{i=1}^N (R_i - \bar{R})^2} \quad \text{Where } \bar{R} \text{ is Bin_Mean_R} \quad (4)$$

Similarly the parameter values Gi and Bi for Green and Blue planes are calculated respectively.

III. FEATURE VECTOR MATCHING AND INDEXING

Once the feature vector databases are prepared, comparison process has to be followed. In this work feature vector is extracted for the query image just the same way it was extracted for all database images. Two similarity measures, Euclidean distance and absolute distance [24], [25] are used in actual comparison process. When user enters the query image; feature vector will be extracted and

then Euclidean and Absolute distance between these two is calculated using following equations 2 and 3 respectively. Now to find the match with query; we have selected the images from the database having less distances in the retrieval set. In this process no thresholding is done but to limit this retrieval, we have taken first 100 images having less distances are selected for the final retrieval (Out of 1000). This may include either relevant or irrelevant images.

Euclidean Distance :

$$D_{QI} = \sqrt{\sum_{i=1}^n |(FQ_i - FI_i)|^2} \quad (5)$$

Absolute Distance:

$$D_{QI} = \sum_{i=1}^n |(FQ_i - FI_i)| \quad (6)$$

IV. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Database and Query Image

All the approaches discussed above are experimented with the database of 1000 bmp images (Augmented Wang Database). It includes 10 different categories of images wherein each of them has got 100 images of their own (Includes Flower, Sunset, Mountain, Building, Bus, Dinosaurs, Elephant, Barbie, Mickey and Horse). Sample images from the database are shown in Figure 2.

In this system query is provided as an example image. For which the feature extraction process will be applied and query image feature vector will be obtained. This will be compared with all database feature vectors using two similarity measure given in Equation 5 and 6. For our experiments we have taken 100 images selected randomly from the database itself to be given as query to this system to tests its performance.



Figure 2. Sample Database Images from 10 Different Classes

4.2 Performance Evaluation

In most of the CBIR system, two parameters are widely used for the performance evaluation that are precision and recall [30],[31]. In our work as we have not chosen the threshold on trial and error which sometimes very tedious for the distance calculated between query and database images. Instead we are taken the initial string of 100 images to be retrieved as the output for the given query. Before we do this it is required that distances should be in sorted order so that first 100 images closer to query will be obtained. As we are taking first 100 for retrieval and we have total 100 images for each class in database this generates the plot of 'Precision and Recall Cross Over Point'. We are using two more new parameters given in Equations 7 and 8 respectively as 'LIRS- Length of Initial Relevant String' and 'LSRR- Length of the String to Retrieve all Relevant images' for the comparison and analysis of the results to evaluate the system performance [33].

$$LIRS = \frac{\text{Length_of_Initial_Relevant_String}}{\text{Total_images_of_that_class_in_database}} \quad (7)$$

$$LSRR = \frac{\text{Length_of_String_to_Retrive_all_Relevant_images}}{\text{Total_images_in_database}} \quad (8)$$

4.3 Results Discussion and Analysis

As discussed above in section 2.3 we have total 9 feature vector databases as 3 for each plane based on three different formats of bins. Results obtained for 100 queries are shown in following tables for each plane and each bin parameter ('Bin_Total_R_G_B', 'Bin_Mean_R_G_B' and 'Bin_Standard_deviation_R_G_B'). Table 1, 2 and 3 shows the results of 100 query images for Red, Green and Blue plane for all types of bins using two similarity measures Euclidean Distance (ED) and Absolute Distance (AD). The values obtained in last row of all the tables are the average of 10 queries executed from each of the 10 classes; last row is the total of all relevant images retrieved out of 10,000.

In these results we can observe that performance of the bins goes on increasing from parameter Bin_Total_R_G_B' to Bin_Standard_deviation_R_G_B'. For Bin_Total_R_G_B' relevant retrieval is around 4000 and 4100 for ED and AD respectively. For Bin_Mean_R_G_B' it is around 4100 and 4300 for ED and AD respectively and it is very good about 4400 and 4800 for ED and AD respectively for 'Bin_Standard_deviation_R_G_B'.

Table1. Retrieval Result of Red_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

R-Plane						
Total_Ret	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	329	384	326	397	270	326
Sunset	418	378	402	411	338	292
Mountain	269	323	317	348	388	415
Building	247	291	355	359	401	427
Bus	440	460	485	506	402	441
Dinosaur	430	414	462	619	798	873
Elephant	391	385	365	401	478	478
Barbie	251	258	297	304	359	334
Mickey	497	505	580	609	542	589
Horses	635	629	576	487	518	581
TOTAL	3907	4027	4165	4441	4494	4756

Table2. Retrieval Result of Green_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

G-Plane						
Total_Ret	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	546	624	609	616	414	451
Sunset	462	492	304	274	416	386
Mountain	257	289	230	280	283	330
Building	235	288	333	335	399	443
Bus	408	421	358	416	380	406
Dinosaur	425	397	486	674	809	866
Elephant	385	397	312	349	412	389
Barbie	247	251	254	285	251	283
Mickey	501	500	534	610	500	534
Horses	497	532	391	399	439	491
TOTAL	3963	4191	3811	4238	4303	4579

If these retrieval results compared in terms of precision and recall cross over point, it has achieved quite good results, 0.4 for Bins_ total and Mean, and 0.5 for Bins standard deviation for the absolute distance. We can see that absolute distance is performing well for all the approaches and proving its best as similarity measure. Performance of these results are evaluated using two parameters explained in equation 7 and 8 and expressed in Tables 4, 5, 6 and 7, 8, 9 for LIRS and LSRR respectively for R, G and B planes

Table3. Retrieval Result of Blue plane, for 100 queries using ED and AD for Bin_Total, Mean and Standard deviation R_G_B'

B-Plane						
Total_Ret	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	330	406	490	513	524	577
Sunset	566	575	458	457	334	441
Mountain	338	328	283	348	259	283
Building	224	273	243	289	284	348
Bus	409	400	471	531	393	424
Dinosaur	406	368	553	642	669	670
Elephant	300	309	256	274	250	262
Barbie	253	251	267	299	312	327
Mickey	506	493	464	571	367	428
Horses	576	596	536	473	494	566
TOTAL	3908	3999	4021	4397	3886	4326

Table5. LIRS Result of Green_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

G-Plane						
LIRS	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	8	13	5	6	9	9
Sunset	6	5	4	3	10	8
Mountain	3	3	2	2	3	3
Building	2	2	2	2	3	3
Bus	10	13	9	6	5	6
Dinosaur	11	12	9	20	29	39
Elephant	4	5	2	3	5	4
Barbie	3	4	4	5	6	9
Mickey	11	13	9	13	11	14
Horses	10	14	9	8	15	14
Percentage %	7	8	6	7	10	11

Table4. LIRS Result of Red_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

R-Plane						
LIRS	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	13	13	5	6	5	8
Sunset	8	7	6	5	5	5
Mountain	2	3	2	3	2	2
Building	2	3	2	2	3	4
Bus	11	16	9	10	4	4
Dinosaur	11	11	10	22	26	45
Elephant	5	5	3	3	8	7
Barbie	5	5	7	6	12	14
Mickey	14	16	10	11	20	23
Horses	13	17	16	13	18	24
Percentage %	9	10	7	8	10	14

Table6. LIRS Result of Blue_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

B-Plane						
LIRS	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	6	6	9	10	17	19
Sunset	9	9	7	9	7	11
Mountain	4	3	2	2	3	3
Building	3	3	3	3	3	3
Bus	12	17	19	15	9	11
Dinosaur	10	11	13	20	16	18
Elephant	3	4	2	4	4	4
Barbie	5	5	6	6	6	8
Mickey	10	8	10	12	14	13
Horses	7	12	8	9	18	20
Percentage %	7	8	8	9	10	11

All the entries shown in table 4, 5 and 6 are actually the average of 10 queries from each class using similarity measures ED and AD. Last row entries in all three tables is average percentage. We can observe in above tables 4, 5 and 6 that the length of initial string of relevant images is good for 'Bin_Standard_deviation_R_G_B' compared to the other two types of bins with absolute distance as a similarity measure.

Table7. LSRR Result of Red_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

R-Plane						
LSRR	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	71	68	66	66	99	99
Sunset	81	83	82	81	99	99
Mountain	68	69	70	70	64	62
Building	78	74	58	60	79	67
Bus	64	56	63	55	83	79
Dinosaur	98	99	75	53	16	14
Elephant	86	89	63	64	63	52
Barbie	99	100	99	98	98	99
Mickey	100	100	96	97	96	81
Horses	56	61	70	74	54	45
Percentage %	80	80	74	72	75	70

Table8. LSRR Result of Green_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

G-Plane						
LSRR	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	52	33	46	56	77	79
Sunset	66	71	79	79	97	99
Mountain	73	77	85	85	82	79
Building	80	79	69	69	75	55
Bus	69	66	63	60	77	79
Dinosaur	98	100	68	48	21	15
Elephant	87	91	68	67	65	68
Barbie	99	100	99	99	99	99
Mickey	100	100	100	99	97	94
Horses	63	70	75	75	88	85
Percentage %	79	78	75	74	78	75

Table9. LSRR Result of Blue_plane, 100 queries using ED and AD for Bin_Total, Mean and Standard_deviation_R_G_B'

B-Plane						
LSRR	RGB_BINS		MEAN_BINS		STD_BINS	
QUERY	ED	AD	ED	AD	ED	AD
Flower	55	53	61	70	61	70
Sunset	39	49	58	61	91	89
Mountain	75	81	92	91	93	95
Building	80	80	76	78	82	85
Bus	71	70	65	65	78	84
Dinosaur	98	100	63	47	33	35
Elephant	87	90	68	70	92	91
Barbie	99	100	99	99	97	100
Mickey	100	100	99	98	99	99
Horses	52	49	66	72	62	65
Percentage %	76	77	75	75	79	81

Table 7, 8 and 9 showing the results of the second evaluation parameter LSRR, which identifies the length of the string of images required to retrieved all the images of query class from the database, When we observed these results of the second parameter LSRR, it can be noticed that to retrieve all relevant which means to make the recall value 1 the length of string of images required to retrieve is almost 70 to 80 %. Overall, when we compared the approaches standard deviation is performing better among three parameters. It requires length around 70 to 75 % to retrieve all relevant images from the database

The results obtained for 3 planes are analyzed and we tried to refine them further using the three criterion [26] , Criterion1, 2 and 3. Using these criteria we could combine the three sets of results obtained for 3 planes separately. Based on the way they are combined they have produce different results.

Criterion1 is called Strong_R_G_B : According to this criterion the image will be retrieved in final set only and only if it being retrieve in all the three planes.

Criterion2: is Called Average _R_G_B: According to this criterion the image will be retrieved in final set if it being retrieve in any of the two planes out of three

Criterion 3 is called Weak R_G_B : According to this the image can be retrieved in final set even if it being retrieve in any one plane out of the three results.

R,G and B Results combined using these three criteria for all bin parameters and for both the similarity measures are plotted in following charts 1 , 2 and 3 for 'Bin_Total_R_G_B', 'Bin_Mean_R_G_B' and 'Bin_Standard_deviation_R_G_B' respectively using both similarity measures ED and AD. The values plotted are the precision recall cross over points obtained for the final refined results. These values are plotted as average percentage retrieval of similarity in terms of cross over points.

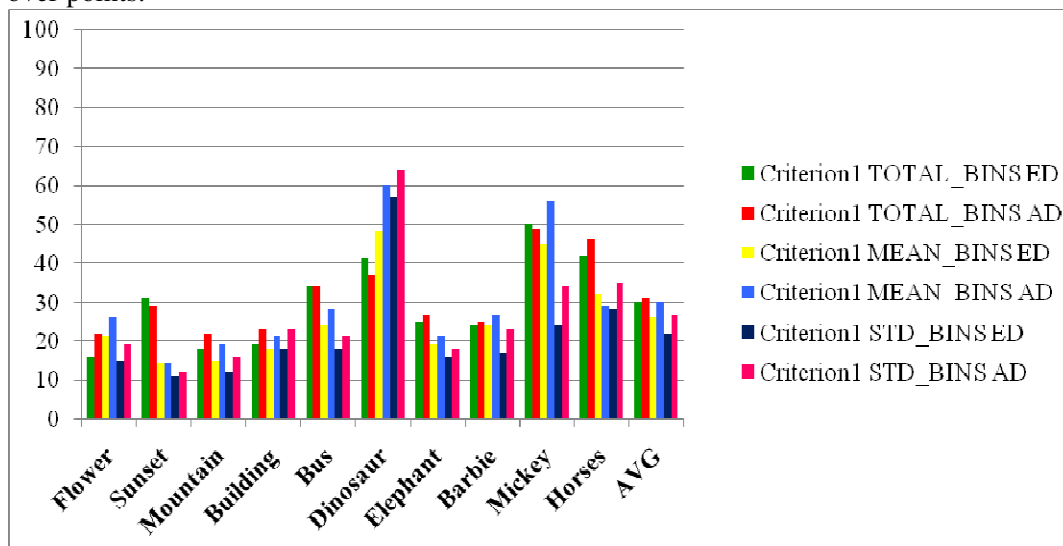


Figure 3: Criterion1 Applied for Total_Bins, Mean_Bins and STD_Bins

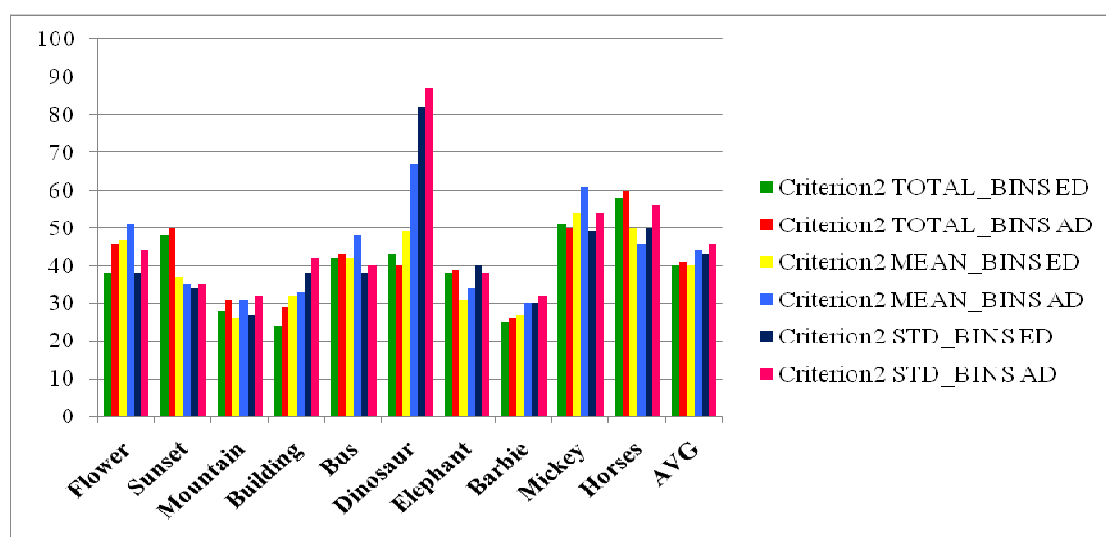


Figure 4: Criterion2 Applied for Total_Bins, Mean_Bins and STD_Bins

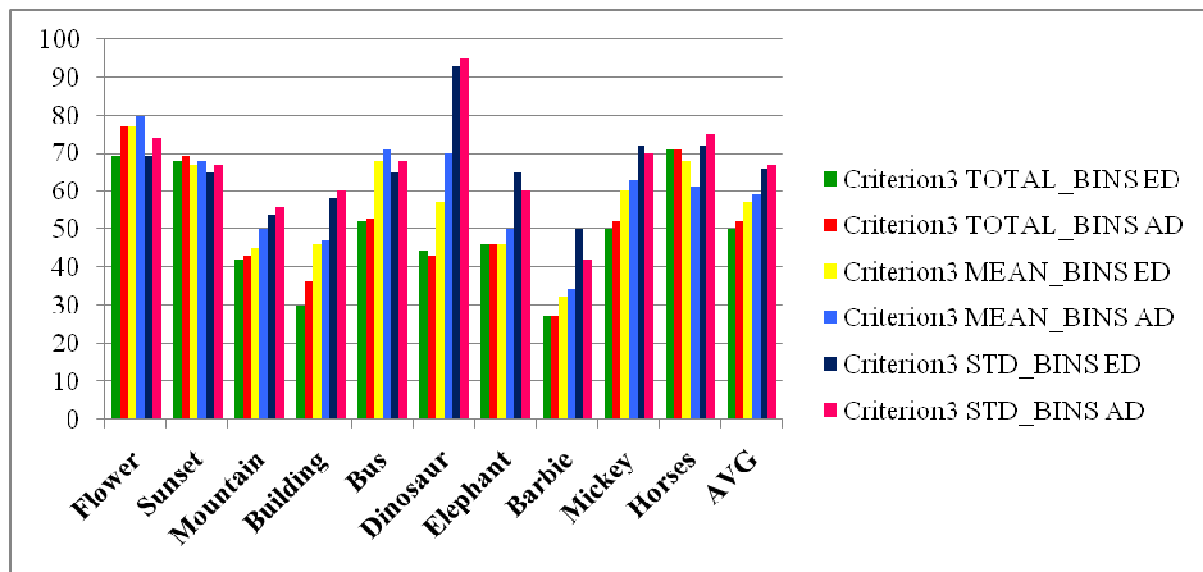


Figure 5: Criterion3 Applied for Total_Bins, Mean_Bins and STD_Bins

V. CONCLUSION

In this work information obtained from unequalized histograms is utilized efficiently and effectively to form the bins which has produced very good results just on the basis of color information of the image. Separating the image into 3 planes and obtaining their histograms separately produces the discriminate color information from one image which is used to represent the feature vector.

Different parameters used to represent the pixel's color information produced different bins formats named Bin_Total_R_G_B bins, Bin_Mean_R_G_B bins and Bin_Standard Deviation_R_G_B bins. Among these three types of feature vector of 27 components, parameter standard deviation is giving very good results as compared to the other two.

Similarity measure selection is highlighting the effective use of Absolute distance for this system. It reduces the computational as well as time complexity and has produced quite good results as compared to the Euclidean distance.

Positive change in the retrieval results is obtained after combining and refining them using three criteria. Observing the figure 3, 4 and 5 it can be delineated that among these results criterion3 is proving its best by giving the best PRCP (Precision Recall Cross over Point) values, which is reached to 9 for few query images and also noticed that it is above 7 for 60% of the query images.

REFERENCES

- [1] J.R.Smith and S.-F Chang, "Integrated spatial and feature image query", *Multimedia Syst.* 7(2), 129-140(1999).
- [2] R. Schettini, G. Ciocca, S Zuffi. "A survey of methods for colour image indexing and retrieval in image databases", *Color Imaging Science: Exploiting Digital Media*, (R. Luo, L. MacDonald eds.), J. Wiley, 2001.
- [3] J. Smith and S. Chang. "Automated image retrieval using color and texture". Technical Report CU/CTR 408-95-14, Columbia University, July 1995.
- [4] John R. Smith and Shih-Fu Chang, "Tools and Techniques for Color Image Retrieval" Columbia University Department of Electrical Engineering and Center for Telecommunications Research New York, N.Y. 10027.
- [5] Y. Liu, D. Zhang, G.Lu, W.Y. Ma., 2007. "A survey of content-based image retrieval with high-level semantics", *The journal of the recognition society*, pp.262- 282 DOI:10.1016/j.patcog.2006.04.045
- [6] P. B. Thawari & N. J. Janwe, "CBIR Based On Color And Texture" *International Journal of Information Technology and Knowledge Management* January-June 2011, Volume 4, No. 1, pp. 129-132.
- [7] M. J. Swain and D. H. Ballard. "Color indexing", *International Journal of Computer Vision*, 7:1 1991.

- [8] N.K.Kamila, ,Pradeep Kumar Mallick, Sasmita Parida B.Das,” Image Retrieval using Equalized Histogram Image Bins Moments”, Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010 for International Conference [ICCT-2010], 3rd-5th December 2010
- [9] H.B.Kekre, Sudeep D. Thepade, “Image Blending in Vista Creation using Kekre's LUV Color Space”, SPIT-IEEE Colloquium and International Conference, Sardar Patel Institute of Technology, Andheri,Mumbai, 04-05 Feb 2008.
- [10] James Hafner, Harpreet S.Sawhney, Will Equits, Myron Flickner and Wayne Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 17, No. 7, July 1995.
- [11] M. Rautiainen and D. Doermann, “Temporal Color Correlograms for Video Retrieval”,Pattern Recognition, Proceedings of 16th International Conference, vol. 1, Aug. 2002, pp. 267-270
- [12] Ryszard S. Chora's"Image Feature Extraction Techniques and Their Applications for CBIR andBiometrics Systems" international journal of biology and biomedical engineering,2007.
- [13] P.S.Suhasini , 2dr. K.Sri Rama Krishna, Dr. I. V. Murali Krishna, “Cbir Using Color Histogram Processing”, Journal of Theoretical and Applied Information Technology, 2005 - 2009 JATIT.
- [14] J. Domke, Y. Aloimonos, "Deformation and viewpoint invariant color histograms", Proc.BMVC (British Machine Vision Conference), September 2006, Edinburgh, UK.
- [15] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih, “Image indexing using color correlograms” InProc. of the IEEE Computer Society Conference on Vision and Pattern Recognition, pages 762–768,1997.
- [16] K. Konstantinidis , A. Gasteratos , I. Andreadis, “Image retrieval based on fuzzy color histogram processing”, Optics Communications 248 (2005) , 375–386.
- [17] S. Flank, P. Martin, A. Balogh, and J. Rothery. PhotoFile: “A digital library for image retrieval”. InProc. of International Conference on Multimedia Computing and Systems, Washington, DC, May1995. IEEE Computer Society
- [18] V. Vijaya Kumar¹, N. Gnaneswara Rao², A.L.Narsimha Rao³, and V.Venkata Krishna⁴“IHBm: Integrated Histogram Bin Matching For Similarity Measures of Color Image Retrieval”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol 2, No.3, September 2009.
- [19] Shengjiu Wang, “A Robust CBIR Approach Using Local Color Histograms” report submitted to University of aberta
- [20] Sangoh Jeong , “Histogram-Based Color Image Retrieval”, Psych221/EE362 Project Report, Mar.15, 2001.
- [21] A Vadivel, A K Majumdar, Shamik Sural, “ Perceptually Smooth Histogram Generation from the HSV Color Space for Content Based Image Retrieval”.
- [22] Ole Andreas, Flaaton Jonsgard, “Improvements on color histogram based CBIR,2005”
- [23] Neetu Sharma, Paresw Rawat, and jaikaran Singh, “Efficient CBIR Using Color Histogram Processing”, Signal & Image Processing : An International Journal(SIPIJ) Vol.2, No.1, March 2011.
- [24] H.B.Kekre, Dharendra Mishra, “Sectorization of DCT-DST Plane for Column wise Transformed Color Images in CBIR” ICTSM-11, at MPSTME 25-27 February, 2011 on Springer Link.
- [25] H. B. Kekre , Kavita Sonawane, “Feature Extraction in Bins Using Global and Local thresholding of Images for CBIR” International Journal Of Computer Applications In Applications In Engineering, Technology And Sciences, ISSN: 0974-3596, October '09 – March '10, Volume 2 : Issue 2 .
- [26] H. B. Kekre , Kavita Sonawane, “Query Based Image Retrieval Using kekre's, DCT and Hybrid wavelet Transform Over 1st and 2nd Moment” International Journal of Computer Applications (0975 – 8887), Volume *– No.*, October 2011.
- [27] H.B.Kekre, Dharendra Mishra “Sectorization of Full Kekre's Wavelet Transform for Feature extraction of Color Images”,(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No.2, February 2011.
- [28] H. B. Kekre, Kavita Sonawane, ‘Retrieval of Images Using DCT and DCT Wavelet Over Image Blocks” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011.
- [29] H. B. Kekre, Kavita Sonawane “Standard Deviation of Mean and Variance of Rows and Columns of Images for CBIR” WASET International Journal of Computer, Information and System Science and Engineering (IJCISSE), Volume 3, Number 1, pp.8-11, 2009.
- [30] Dr. H. B. Kekre Kavita Sonavane “CBIR Using Kekre's Transform over Row column Mean and Variance Vector ” (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05,2010, 1609-1614.
- [31] Dr. H. B. Kekre, Dharendra Mishra “Image Retrieval using DST and DST Wavelet Sectorization”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.

- [32] Dr. H. B. Kekre, V.A. Bharadi, S.D. Thepade , B.K. Mishra, S.E. Ghosalkar , S.M. Sawan “Content Based Image Retrieval Using Fusion of Gabor Magnitude and Modified Block Truncation Coding” International Conference on Emerging Trends in Engineering & Technology (ICETET_10) 19-21 November 2010, held at BITS Goa, by G.H. Raison College of Engineering, Nagpur, India. IEEE Xplore.
- [33] H B Kekre, D. Mishra “Full DST Sectorization for Feature Vector Generation in Content Based Image Retrieval” pp.307-314, ICWET-2011, February 26-27, 2011 Uploaded on ACM Portal.

Authors Biographies

H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S. Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. For last 13 years worked as a Professor in Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. He is currently Senior Professor working with Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Vile Parle(w), Mumbai, INDIA. He has guided 17 Ph.D.s, 150 M.E./M.Tech Projects and several B.E./B.Tech Projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networks. He has more than 450 papers in National / International Conferences / Journals to his credit. Recently twelve students working under his guidance have received best paper awards. Five of his students have been awarded Ph. D. of NMIMS University. Currently he is guiding eight Ph.D. students. He is member of ISTE and IETE.



Kavita V. Sonawane has received M.E (Computer Engineering) degree from Mumbai University in 2008, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engg, SVKM's NMIMS University, Vile-Parle (w), Mumbai, INDIA. She has more than 8 years of experience in teaching. Currently working as a Assistant professor in Department of Computer Engineering at St. Francis Institute of Technology Mumbai. Her area of interest is Image Processing, Data structures and Computer Architecture. She has 7 papers in National/ International conferences / Journals to her credit. She is member of ISTE.



ADAPTIVE ALGORITHM FOR CALIBRATION OF ARRAY COEFFICIENTS

K. Ch. Sri Kavya¹, B. V. Raj Gopala Rao², Gopala Krishna.N², J. Supriyanka³,
J.V.Suresh², Kota Kumar², Habibulla Khan⁴, Fazal Noor Basha⁵

¹Assoc. Professor, Dept of ECE, K.L.University, Vaddeswaram, A.P., India

²PG Student, Dept of ECE, K.L.University, Vaddeswaram, A.P., India

³Asst. Professor, Dept of EIE, Gitam University, Visakhapatnam, A.P., India

⁴Professor, Dept of ECE, K.L.University, Vaddeswaram, A.P., India

⁵Asst. Professor, Dept of ECE, K.L.University, Vaddeswaram, A.P., India

ABSTRACT

Phased array antennas are deployed in electronic systems where high beam directivity and/or electronic beam scanning is desired in applications ranging from radar systems to smart antennas in wireless communication where errors such as random and/or correlated fluctuations present in the excitation coefficients of a phased array can degrade its performance. The errors due to random environmental changes, mechanical variations, assembly inaccuracies, mutual coupling effects and mistune or failure of amplifiers and phase shifters etc may cause undesirable effects such as decrease in directivity, increase in side lobes and steering the beam in wrong direction. In this paper, an adaptive algorithm for the excitation of coefficients of the phased array is demonstrated. Here a linear array is considered and the knowledge of the reference signal generated by the desired array at the near-field sensors is assumed and the fluctuations in the coefficients of the actual array are corrected by means of a gradient based least mean square adaptive algorithm. Requirements for the algorithm to converge, its performance without receiver noise and the effects of the dither parameters are studied.

KEYWORDS: Dithering, Antenna array, phased array, least mean square & near field sensors.

I. INTRODUCTION

Phased array antennas [1] have acquired a great demand in many applications ranging –from radar systems to smart antenna systems where high beam directivity and electronic scanning of beam in specific and/or different directions is desired. Within the visible or infrared spectrum of electromagnetic waves it is also possible to construct optical phased arrays. They are used in wavelength multiplexers and filters for telecommunication purposes, laser beam steering, and holography. Synthetic array heterodyne detection is an efficient method for multiplexing an entire phased array onto a single element photo detector.

Irrespective of high costs and complex structure of phased array antennas, there are many advantages which include low probability of interception, high jamming resistance, multifunction operation by emitting several beams simultaneously, ability to permit the beam to jump from one target to next in a few milliseconds etc.

But it has well known that random and fluctuation errors due to correlation present in the excitation coefficients (amplitude, phase) of phased arrays may degrade its radiation pattern that affects the desired characteristics. The desired characteristics include decrease in directivity, increase in side lobes and beam steering in wrong direction.

This degradation in desired characteristics may particularly be severe in applications of high performance arrays such as in satellite communication where high directivity and low side lobes are often required. Degradation in radiation pattern requires high transmitting power or causing

interference to neighboring satellites. The sources of errors include those introduced by the random variations in the radio propagation channel, mechanical variations of the phased array set up, inaccuracies caused during assembly, mutual coupling effects between neighboring elements in the array etc.

The methodology adopted in this paper is given based on the linear array with two near field sensors where the coefficients are considered to be perturbed due to the environmental effects around the antenna. Later they are dithered and then corrected using the adaptive algorithm of least mean square technique.

The results are presented for a broad side Taylor array of 32 elements with side lobe level of -25 dB. The total length of the array is assumed to be $2L = 15.5\lambda$ with element spacing of 0.5λ . The normalized current distribution of true and actual array with respect to the element number plotted and for the purpose of illustration, we perturb the true coefficients randomly with the magnitude varied on a dB scale using log-normal distribution with an RMS deviation of 2 dB and the phase varied uniformly with an RMS deviation of 10° . The current distribution of perturbed true and actual array along with the dithering applied is also plotted. Later the near field sensed magnetic field and far field magnetic fields are plotted showing the degradation in the sidelobe level and broadened beamwidth. Then using the adaptive algorithm the performance is improved and results explain themselves that after the correction the performance is almost close that of the true array.

II. METHODOLOGY

2.1. Coefficients in Linear Array

Here a linear array [5], [6], [7] was considered with dipole elements arranged along x -axis with spacing between the elements as ' d ' as shown in Fig 1. The axis of the dipoles are assumed to lie along the z -axis the total number of elements are assumed as ' N '. The distance of the observation point from the n -th element is denoted by ' R_n '. The total length of the array is given by $2L = (N-1)d$. Two near field sensors are placed towards 1st element and N th element. The near field sensors are assumed to sample the magnetic field, although the theory developed is equally valid for an electric field sensor.

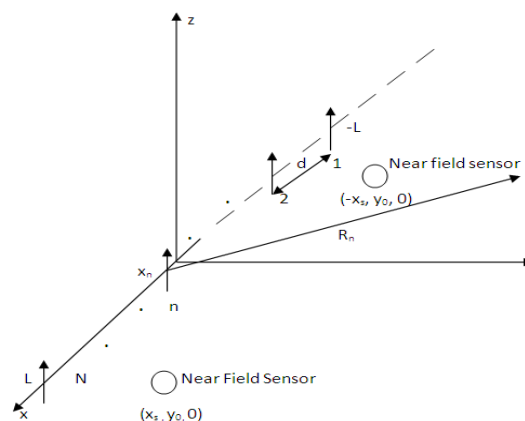


Fig 1: A Uniform Linear Array comprised of dipoles with proposed positions of near field sensors.

In the following, the array coefficients with C_n is treated as true array (or desired array) and that with \tilde{C}_n as actual array. The normalized complex current excitation coefficient of the n -th element is denoted by $C_n = a_n e^{i\psi_n}$, where a_n and ψ_n are the magnitude and phase respectively. Now for automatically correcting the coefficients \tilde{C}_n ,

We introduce dithering [10][11][12][13] and [14] which means introducing pseudo random fluctuations into magnitude and phase of the coefficients for both true and actual array. Here we assume a log-normal distribution with a standard deviation of σ dB for the magnitude and a uniform distribution with a maximum deviation of Δ for the phase. The dithered magnitudes and phases of the true array are

$$\tilde{a}_n = a_n e^{\alpha v_n}, \quad \alpha = 0.05 \ln(10) \sigma$$

$$\tilde{\psi}n = \psi n + \mu n \Delta$$

Where v_n is a unit-variance, zero mean Gaussian random variable and μ_n is a uniform random variable. Now, the dithered magnetic field vector is computed using the complex current distribution of true dithered and actual dithered coefficients. These dithered magnetic fields due to the true and actual array are assumed to be observed at the near field sensors. As two near field sensors are placed near the first and the N-th, vector addition of the magnetic fields is considered. Finally, the error signal which is obtained by subtracting the true dithered coefficients from actual dithered coefficients is minimized using a gradient based adaptive algorithm that can be devised to nullify unwanted deviations.

2.2. Adaptive algorithm

An adaptive algorithm [2][3][4][8] and [9] is an algorithm that varies the weights of the phased array based on the received data in order to improve the signal strength and reduce the bit error rate. The algorithm is crucial in steering the main beam of the antenna array. Different algorithms have different characteristics, i.e. different convergence rates, computation complexity, and effectiveness. Least mean square algorithm (LMS) was used in the present work due to its simplicity as it does not require correlation function calculation nor it does require matrix inversions. The block diagram of LMS algorithm for adaptive beam forming is shown in Fig 3.

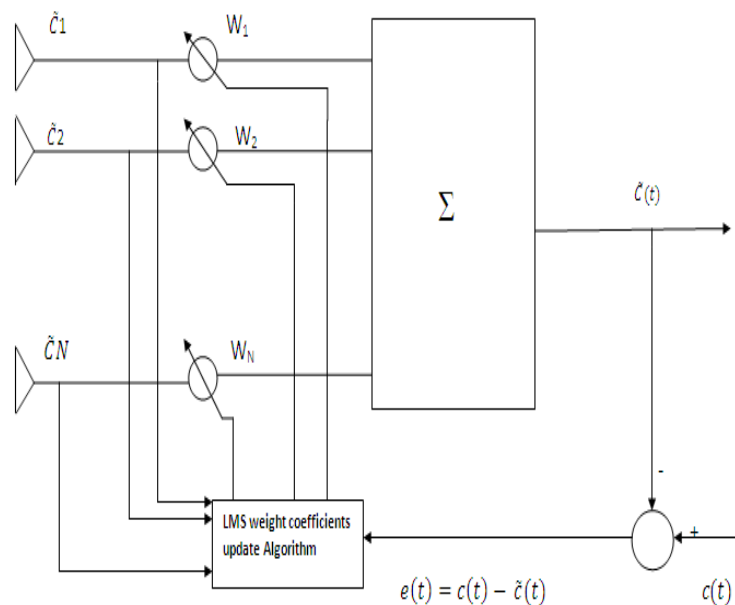


Fig 2: LMS adaptive beam forming network

From the above Fig 2, the actual dithered coefficients sensed by the near field sensors are scaled using corresponding weights computed by the least mean square (LMS) update algorithm based on minimum mean square (MSE) criterion. These outputs are linearly combined such that the error is minimized. The error signal which is obtained by subtracting the linearly combined signal from the reference signal i.e. the true dithered signal is fed to the LMS update algorithm which does successive corrections to the weight vector by iterative process and eventually leads to the minimum value of the mean squared error.

The LMS algorithm initiated with some arbitrary value for the weight vector is seen to converge and stay stable for

$$0 < \mu < 1/\lambda_{\max}$$

Where μ is step size which is chosen very small such that the algorithm converges very slowly and may be more stable around the minimum value.

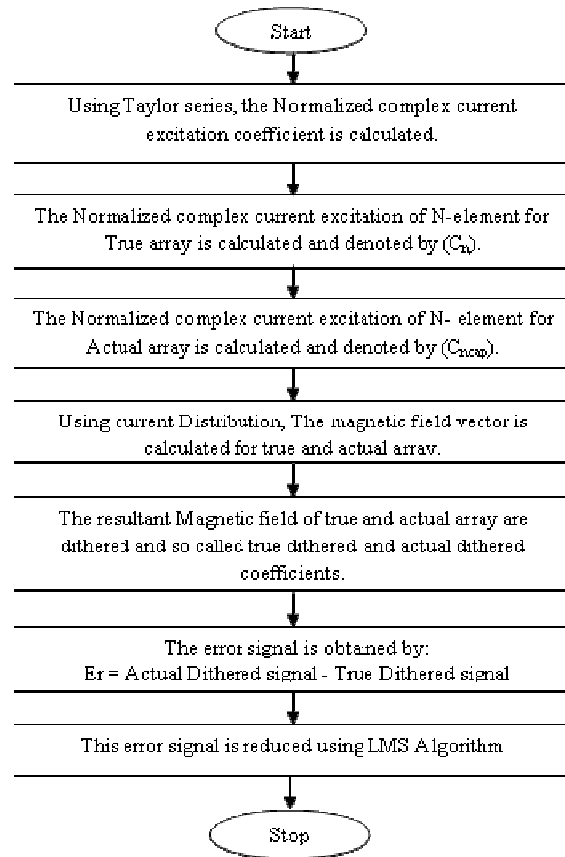


Fig 3: Flow Chart of proposed technique.

III. SIMULATION RESULTS

Results are presented below for a broad side Taylor array of 32 elements with side lobe level of -25 dB. The total length of the array is assumed to be $2L = 15.5\lambda$ with element spacing of 0.5λ . The simulations are all performed using programming in MATLAB Software. These simulated MATLAB graphs made our analysis easier.

The normalized current distribution of true and actual array with respect to the element number is shown in Fig 4 and for the purpose of illustration, we perturb the true coefficients randomly with the magnitude varied on a dB scale using log-normal distribution with an RMS deviation of 2 dB and the phase varied uniformly with an RMS deviation of 10^0 . The current distribution of perturbed true and actual array are shown in Fig 5.

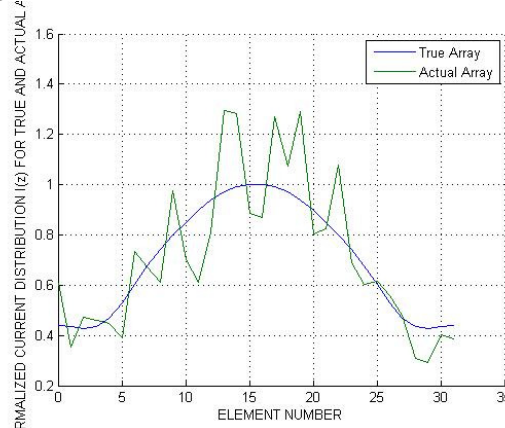


Fig 4: Normalized current distribution of True and Actual perturbed Array.

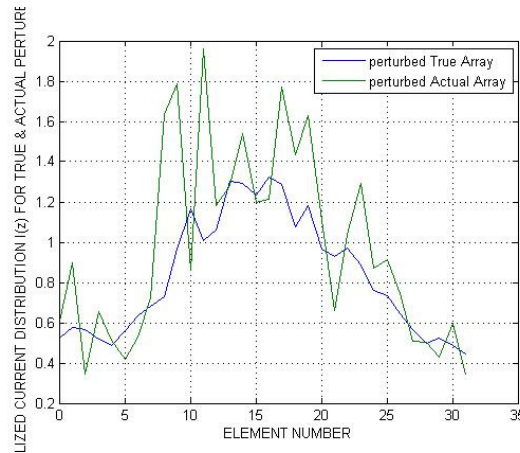


Fig 5: Normalized current distribution of True and Actual Array.

The far-zone magnetic field strength for the true and actual perturbed array as a function of lateral displacement x for $y = 10R_f$ and $z=0$ is shown in Fig 6. As a result of fluctuations introduced, the side lobes have increased substantially and the main lobe slightly broadened.

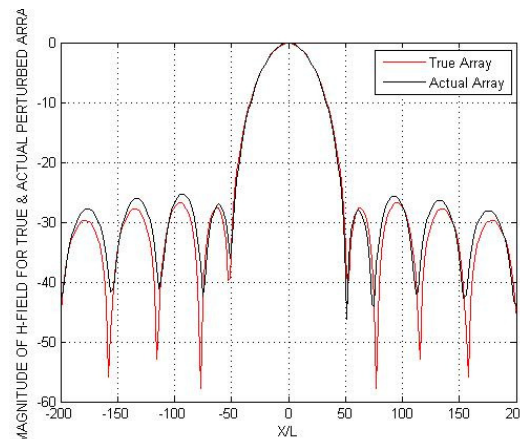


Fig 6: Perturbed True and Actual Magnetic fields of a Linear Array.

A near -field sensor is assumed to be located in the $z = 0$ plane at $x = x_s = 1.1L$ and $y = y_0 = R_f/100 = 4.805\lambda$ as shown in Fig 1. The true and actual coefficients are dithered using an RMS deviation of magnitude with 3dB and that of RMS deviation of phase with 12° . The normalized current distribution of dithered true and actual array are shown in Fig 7 and the dithered true and actual near fields are shown in Fig 8.

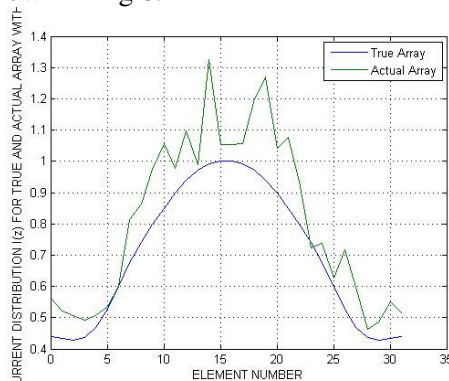


Fig 7: Normalized current distribution of dithered true and actual array.

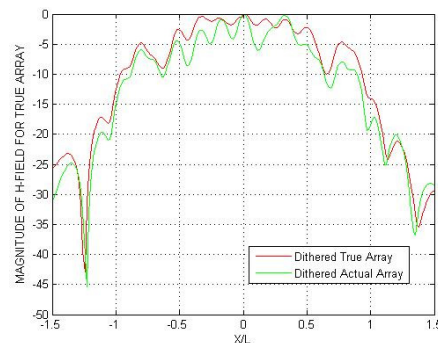


Fig 8: Fields at near field sensor 1 for dithered true and actual array.

Another near -field sensor is assumed to be located in the $z = 0$ plane at $x = x_s = -1.1L$ and $y = y_0 = R_f/100 = 4.805\lambda$ as shown in Fig 1. The true and actual coefficients are dithered using an RMS deviation of magnitude with 3dB and that of RMS deviation of phase with 12° . The dithered true and actual near fields are shown in Fig 8 and the vector addition of the fields those sensed by the proposed near field sensors for dithered true and actual array are shown in Fig 9. The Fig 10 shows the application of adaptive correction algorithm applied to the dithered curves of Fig 9. One can notice the approximation of the corrected curve towards the original.

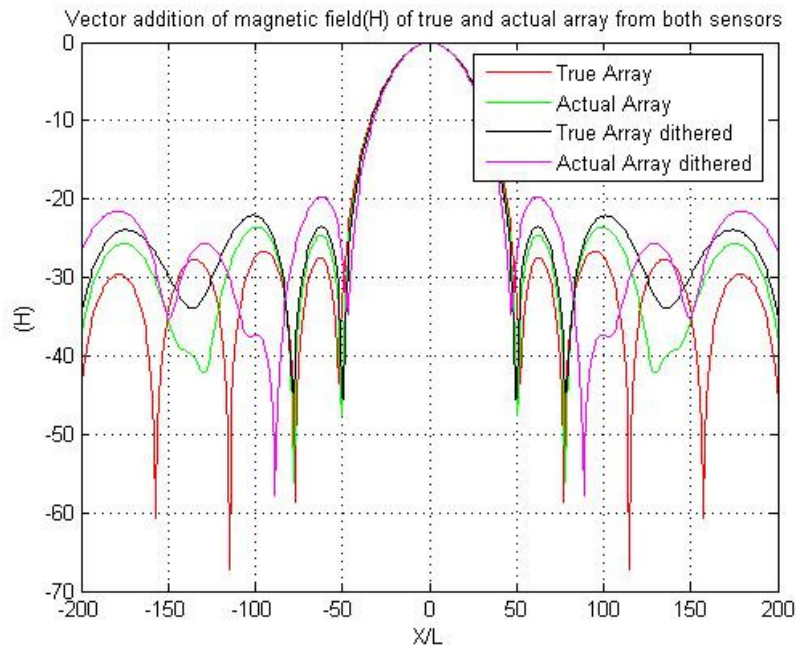


Fig 9: Vector addition of the fields sensed by the proposed sensors for dithered true and actual array

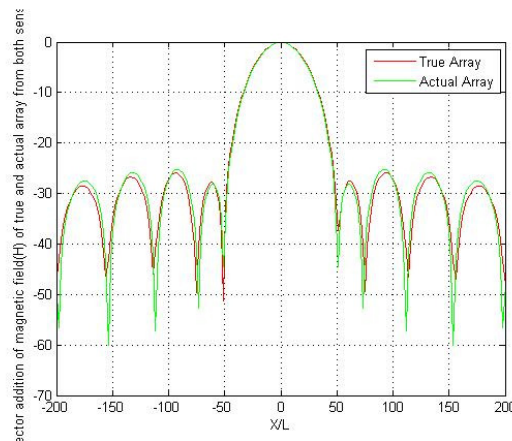


Fig 10: Corrected Field pattern using adaptive algorithm.

IV. CONCLUSIONS

A least mean square type algorithm was presented for correcting the perturbed array coefficients in noise free environment. The robustness of the algorithm has been demonstrated by considering a low side lobe (-25 dB) broad side array with RMS magnitude fluctuations and RMS phase fluctuations in the array coefficients. An adaptive algorithm for automatically correcting the desired excitation coefficients of an antenna array by dithering the coefficients and observing its field in the near zone

by using two near field sensors has been proposed and demonstrated by considering a uniform linear array.

By applying the dithering and correcting algorithm the excitation coefficients are corrected near to their true values. These coefficients suffered from errors earlier and are rectified using our algorithm. By using this algorithm, the performance of the antenna remains close to the specifications even in an environment instead of degraded.

In the future work, adaptive correction of array coefficients by considering the case in noisy environments and also with mutual coupling effects will be analyzed. Further extension can be done with the analysis on Electric fields of the array.

ACKNOWLEDGEMENTS

The authors would like to thank the management of KL University and the faculty of Department of Electronics & Communication Engineering for their constant support.

REFERENCES

- [1] Constantine. A. Balanis, “*Antenna Theory*”, Newyork: John Wiley and Sons.
- [2] Ramakrishna Janaswamy , Dev V. Gupta, Daniel H. Schaubert , “*Adaptive Correction to Array Coefficients through Dithering and Near field Sensing*”, IEEE Trans Antenna Prop, Vol.58, NO.11, 2010.
- [3] Hubregt J. Vesser, “*Array and Phased Array Antenna Basics*”, Newyork : John Wiley and Sons.
- [4] A. Zaknich, “*Principles of Adaptive Filters and Self Learning Systems*”, Springer.
- [5] K.Ch. Sri Kavya, N. Susmitha, K. Priyanka and Dr. N.N. Sastry, “Broadband Phased Arrays in Ku Band for Airborne Communications”, International conference (conducted by IEEE and IETE) ISM-08, Bangalore.
- [6] K.Ch. Sri Kavya, S. Sri Jayalakshmi, Dr. Habibulla Khan, “Hyper beam forming technique using array antenna”, International Journal of systems and technology - Vol. 3 No.2, 2010
- [7] Dennis J. Gaushell, “Synthesis of linear antenna arrays using z-transforms,” IEEE trans on antenna and propagation, 1971 (1).
- [8] L. G. Roberts, “Picture coding using pseudo-random noise,” *Ire Trans. Info. Theory*, vol. 8, pp. 145–154, Feb. 1962.
- [9] K.Ch. Sri Kavya, Dr. Habibulla Khan, “Ku band circular patch array antenna”, International Journal of systems and technology - Vol. 3 No.2, 2010
- [10] I. Schuchman, “Dither Signals And Their Effect On Quantization Noise,” *IEEE trans. Commun. Technology*, pp. 162–165, dec. 1964.
- [11] G. Zames and N. A. Shneydor, “Dither In Nonlinear Systems,” *IEEE trans. Automat Control*, vol. Ac-21, pp. 660–667, oct. 1976.
- [12] T.T. Taylor , “One Parameter Family Of Line Sources Producing Modified $\frac{\sin(\pi u)}{\pi u}$ Patterns,” *Hughes Aircraft co. tech .mem .324, culver city, calif ., contract af 19(604)-262-f-14, september 4, 1953.*
- [13] Taylor, T. T., “Design Of Line-Source Antennas For Narrow Beam Width And Low Side Lobes,” *IRE transactions on antennas and propagation, January, 1955, pp. 16-28.*
- [14] Taylor, T. T., “Design Of Circular Aperture For Narrow Beam Width And Low Side Lobes,” *IRE Transactions On Antennas And Propagation, January, 1960, pp. 17-22.*

AUTHORS BIOGRAPHY

K. Ch. Sri Kavya completed B-Tech in 2003 and got her M-Tech in June 2008 from Acharya Nagarjuna University. She has been working as an Associate Professor at K.L. University since 2004. She has been involved in 5 R&D out sourcing projects from DST, UGC, DRDO, & ISRO. She has been pursuing her PhD in KL University in the area of Antennas and is also interested in the areas of Radars & Communication systems.



Raj Gopala Rao.B.V received his bachelor degree in Electronics and Communication engineering from Pydah College of Engineering, Visakhapatnam in 2007. He worked as project associate in N.S.T.L in the area of underwater sensors and inertial navigation of underwater robotics. Present he is pursuing his M.Tech –Communication and Radar systems in KL University. His interested areas are Inertial Navigation, Underwater Acoustics, Microwaves, Antennas and Radar.



J. Supriyanka received her B.Tech degree in Electronics and Instrumentation engineering from the Pondicherry University, India, in 2006 and the M.E. degree in Electronic Instrumentation engineering from the Andhra University, Visakhapatnam, Andhra Pradesh, India in 2008 respectively. Currently, she is working as an ASSISTANT PROFESSOR in the Department of Electronics and Instrumentation Engineering, Gitam Institute of Technology, Gitam University, Visakhapatnam. Her research interests include MEMS, Smart Sensors, signal processing. She has life time membership in the professional body namely Instrument Society of India (ISoI).



Gopala Krishna.N was born in Krishna dist. In the year 1989. His completed his B.Tech in Sri Saradhi Institute of Engineering and Technology. Presently he is pursuing his M.Tech in Communication and Radar Systems in K.L.University, Guntur. His interested areas are wireless sensor networks, antennas and OFDM.



K. V. V. Kumar was born in 1987 at Guntur district of Andhra Pradesh state, India. He Graduated in Electronics and Communication Engineering from VYCET, JNTU, Kakinada. Presently he is pursuing his M.Tech –Communication and Radar systems in KL University. His interested areas are Image Processing, Antennas, and Communication.



Joga Venkata Suresh was born in Vizianagaram, Andhra Pradesh, in the year 1988. He completed his B.Tech from St. Theresa Institute of Engineering & Technology, affiliated by JNTU Kakinada, in 2009. Now he is pursuing his M.Tech in Communication & Radar systems from K.L.University. His interested areas are antennas and wireless communications.



Habibulla Khan born in India, 1962. He obtained his B.E. from V R Siddhartha Engineering College, Vijayawada during 1980-84. M.E from C.I.T, Coimbatore during 1985-87 and PhD from Andhra University in the area of antennas in the year 2007. He is having more than 20 years of teaching experience and having more than 20 international, national journals/conference papers in his credit. Prof. Habibulla Khan presently working as Head of the ECE department at K.L.University. He is a fellow of I.E.T.E, Member IE and other bodies like ISTE. His research interested areas includes Antenna system designing, microwave engineering, Electro magnetics and RF system designing.



Fazal Noorbasha was born on 29th April 1982. He received his, B.Sc. degree in Electronics Sciences from BCAS College, Bapatla, Guntur, A.P., Affiliated to the Acharya Nagarjuna University, Guntur, Andhra Pradesh, India, in 2003, M.Sc. degree in Electronics Sciences from the Dr. HariSingh Gour University, Sagar, Madhya Pradesh, India, in 2006, M.Tech. Degree in VLSI Technology, from the North Maharashtra University, Jalgaon, Maharashtra, INDIA in 2008, and Ph.D. (VLSI) from Department Of Physics and Electronics, Dr. HariSingh Gour Central University, Sagar, Madhya Pradesh, India, in 2011. Presently he is working as a Assistant Professor, Department of Electronics and Communication Engineering, KL University, Guntur, Andhra Pradesh, India, where he has been engaged in the teaching, research and development of Low-power, High-speed CMOS VLSI SoC, Memory Processors LSI's, Digital Image Processing, Embedded Systems and Nanotechnology. He is a Scientific and Technical Committee & Editorial Review Board member in Engineering and Applied Sciences of World Academy of Science Engineering and Technology (WASET) and Member of International Association of Engineers (IAENG). He has published over 20 Science and Technical papers in various International and National reputed journals and conferences.



SYNTHESIS AND CHARACTERIZATION OF CATALYSTS CONTAINING MOLYBDENUM AND TUNGSTEN AND THEIR APPLICATION IN PARAFFIN ISOMERIZATION

Aoudjit Farid

Faculty of Hydrocarbons and Chemistry; University of Boumerdes

ABSTRACT

The development of new catalysts for normal alkanes isomerization reveals a big importance. Three series of catalysts were prepared by using sol-gel method, using different metal concentration and in two different pH (acid and basic). The first series contains molybdenum, the second contains tungsten and the third contains the two metals (molybdenum and tungsten). The catalysts were characterized by several techniques; the obtained results show that the catalysts prepared in basic pH have a surfaces areas and porous volumes higher than those prepared in acid pH. In order to evaluate the catalytic performances of prepared catalysts, we carried out the catalytic tests with light naphtha (C5, C6) isomerization. The obtained results show that the best conversions are obtained with catalysts prepared into pH basic. The octane numbers of the obtained products (isomerate) are high compared to naphtha. The addition of the obtained products (isomerate) into the pool gasoline allows on the one hand, to obtain gasolines with high octane numbers which respect the environmental standards, and on the other hand, the possibility of developing the light fractions (C5 - C6) resulting from the crude oil distillation unit and condensed split distillation.

KEYWORDS: Isomerization, Pollution, Catalyst, Molybdenum, Tungsten

I. INTRODUCTION

The deterioration of the natural environment by an increasing pollution due to vehicles emissions has been subject to increasingly severe environmental regulations aiming in reducing vehicles emissions by introducing new specifications for fuels.

Among these specifications the total elimination of lead additions to fuels, the reduction of aromatic, olefins and sulphur content, and the increase of octane number.

Environmental reasons are motivating the production of transportation fuels with increase amount of high octane number branched alkanes [1]. These compounds are made by alkylation and by isomerization of straight-chain alkanes. The drawback of these two processes is that they employ harmful and potentially polluting catalysts, such as H₂SO₄, and HF [2, 3].

The isomerization for normal paraffin to isoparaffin is regarded as one of the important technologies for RON improvement in oil refining [4, 5,6] and it is one of the largely used petrochemical processes. It is well known that low temperatures are preferred for alkane isomerization processes because the equilibria favors branched (high-octane-number) products, thus in order to achieve maximum isomer yields, the isomerization of n-paraffins must be carried out at the lowest possible temperatures over highly efficient catalysts [7, 8, 9, 10].

Industrially the catalyst used for alkane's isomerisation is chlorinated platinum-alumina; this catalyst is very sensitive to the poisons and present corrosion problems [11], of this fact the development of new isomerization catalysts reveals a big importance and extensive research has been done in the search for such catalysts, in this context H.Al-Kandari and al [12] have studied the isomerization reaction of n-hexane using molybdenum trioxide partially reduced and deposited on TiO₂, depending

on the weight of molybdenum. These authors reported that at temperature of 623 °C and hydrogen flow rate of 3 to 15 bars the conversion is about 80% with a selectivity of 90% in mono- and di-branched isomers.

Holl and Al [13] showed that the catalysts $\text{MoO}_3 / \text{Al}_2\text{O}_3$, in which the molybdenum is reduced to the metallic state (at 950 °C) catalyzes the isomerization reaction of hexanes.

Hino and Arata [14] have reported that the catalysts Mo/ZnO_2 and W/ZnO_2 prepared by impregnation and calcined between 800 °C and 900 °C are active in the isomerization reactions of n-pentane and n-hexane with a significant participation of the cracking reactions.

Inglesia and Coll [15, 16] show that the tungsten carbides treated with oxygen are highly selective for the isomerization of hexanes and heptanes. They offer a classical bifunctional mechanism,

The sol-gel is one of the catalysts preparation methods; this way of synthesis is a homogeneous process resulting in a continuous transformation of a solution into hydrated solid precursor (hydrogel) [17, 18, 19]. The gelation route (sol-gel process) has several promising advantages over the conventional techniques. It offers better control of the texture, composition, homogeneity and structural properties of the final solids [20, 21, 22, 23, 24, 25].

Our work in this article is to develop new catalysts with molybdenum and tungsten for the alkanes' isomerisation reaction using the sol-gel method.

We developed three series of catalysts, the first series with molybdenum ($\text{MoO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$), the second with tungsten ($\text{WO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$) and the third with the two metals (Mo and W) ($\text{MoO}_3 - \text{WO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$).

Our first objective in this work is to study the metal adding effect (Mo, W) on the catalysts characteristics (surface area, porous volume) and on the alkanes' isomerisation.

The second objective is to study the pH effect on the catalysts characteristics; we prepared catalysts with pH 4.55 and catalysts with pH 7.

The third objective is to study the effect of the metal's nature on the catalysts characteristics; we prepared catalysts with molybdenum, with tungsten and with the two metals (Mo and W).

The obtained catalysts were tested on the isomerisation reaction of light naphtha (C5, C6) resulting from the atmospheric crude oil distillation.

The obtained products (isomerisats) were analyzed and the obtained results shows a considerable increase of octane number and isoparaffins content compared to naphtha, these results confirm the isomerisation properties of the prepared catalysts.

II. EXPERIMENTAL

2.1. Catalysts Preparation

By using sol-gel method three series of catalysts were prepared, the first series contains molybdenum ($\text{MoO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$), the second contains tungsten ($\text{WO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$) and the third contains the two metals (molybdenum and tungsten) ($\text{MoO}_3 - \text{WO}_3 / \text{Al}_2\text{O}_3 - \text{SiO}_2$).

The catalysts were synthesised with different metal concentration (Table 1 and 2) and in two different pH, one acid (pH = 4.5) and the other basic (pH = 10.5).

For the first series of preparation we prepared the acidic gel by mixing the required amounts of aluminum sulphate ($\text{Al}_2(\text{SO}_4)_3 \cdot 18 \text{H}_2\text{O}$) with ammonium heptamolybdate $[(\text{NH}_4)_6 \text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}]$ preliminarily acidified by nitric acid. To the obtained sol, under vigorous stirring, an aqueous solution of sodium silicate (Na_2SiO_3) with $\text{Na}_2\text{O} / \text{SiO}_2 = 2$ was added. The obtained gel have a pH = 4.5. Whereas to obtain a basic gel an aluminum sulphate and ammonium heptamolybdate were added to the sodium silicate solution, the obtained gel have a pH = 10.5.

Then the obtained sol was activated in thermostat at temperature ranging between 60 ÷ 80 °C. This operation consists to exchange undesirable ions, such as Na^+ by ions ammonium, by treating the prepared gel with ammonia sulfate $(\text{NH}_4)_2\text{SO}_4$ at 20 g/l. After the catalyst were washed with water at temperature ranging between 60 ÷ 80 °C. The complete elimination of undesirable ions was tested with AgNO_3 and BaCl_2 solutions. Finally the catalysts were dried at 120 °C for 6 h and calcined at 700 °C.

For the preparation of series 2 and 3 is the same method with series 1, while in serie 2 the used solution is sodium tungstate ($\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$) and in serie 3 the two solutions $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ and $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$ were added.

Table 1. Chemical Composition of Catalysts Prepared at pH = 4.5

Catalyst	% MoO_3	% WO_3	$\text{SiO}_2/\text{Al}_2\text{O}_3$
$(\text{MoO}_3)_{1B}$	4	-	2
$(\text{MoO}_3)_{2B}$	6	-	2
$(\text{MoO}_3)_{3B}$	8	-	2
$(\text{MoO}_3)_{4B}$	12	-	2
$(\text{WO}_3)_{1B}$	-	4	2
$(\text{WO}_3)_{2B}$	-	6	2
$(\text{WO}_3)_{3B}$	-	8	2
$(\text{WO}_3)_{4B}$	-	12	2
$(\text{MoO}_3 - \text{WO}_3)_{1B}$	2	2	2
$(\text{MoO}_3 - \text{WO}_3)_{2B}$	2	6	2
$(\text{MoO}_3 - \text{WO}_3)_{3B}$	6	2	2
$(\text{MoO}_3 - \text{WO}_3)_{4B}$	6	6	2

With: $(\text{MoO}_3)_{iA}$:

i = catalyst number

i = 1, 2, 3, 4

A = acid

Table 2. Chemical Composition of Catalysts Prepared at pH = 10.5

Catalyst	% MoO_3	% WO_3	$\text{SiO}_2/\text{Al}_2\text{O}_3$
$(\text{MoO}_3)_{1A}$	4	-	2
$(\text{MoO}_3)_{2A}$	6	-	2
$(\text{MoO}_3)_{3A}$	8	-	2
$(\text{MoO}_3)_{4A}$	12	-	2
$(\text{WO}_3)_{1A}$	-	4	2
$(\text{WO}_3)_{2A}$	-	6	2
$(\text{WO}_3)_{3A}$	-	8	2
$(\text{WO}_3)_{4A}$	-	12	2
$(\text{MoO}_3 - \text{WO}_3)_{1A}$	2	2	2
$(\text{MoO}_3 - \text{WO}_3)_{2A}$	2	6	2
$(\text{MoO}_3 - \text{WO}_3)_{3A}$	6	2	2
$(\text{MoO}_3 - \text{WO}_3)_{4A}$	6	6	2

With: $(\text{MoO}_3)_{iB}$:

i = catalyst number

i = 1, 2, 3, 4

B = Basic

2.2 Catalysts characterization

2.2.1 Infrared spectroscopy

The prepared catalysts were analyzed by infrared spectroscopy, the obtained spectres shows that the spectres takes the same general form and the bands observed are represented in table 3.

Fig. 1 and table 3 gives the infrared specters of catalysts

Table 3. Infrared Spectres

Bande d'absorption (Cm^{-1})	Caractérisation
3400 – 3500	Attribute to the hydroxyls O-H
1630 – 1620	Attribute to the water molecule vibration.
1060 – 1050	Attribute to the Si-O-Al and Al-O-OH
795 - 790	Elongation of Si-O
500 – 450	Attribute to the metallic oxides

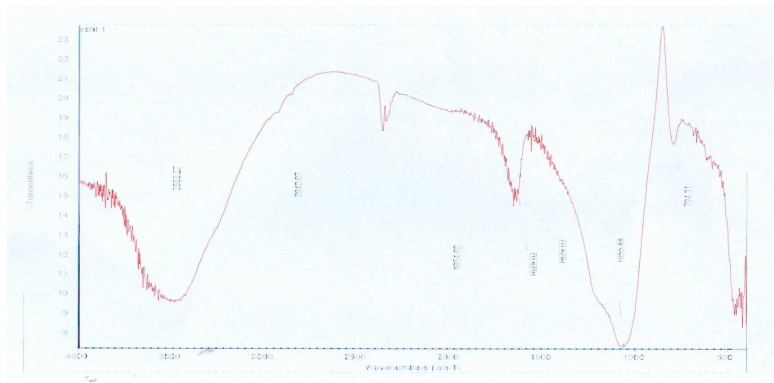


Figure 1. Infrared Spectres

2.2.2 Surface area

The surface area was determined by the BET method; we measured the adsorption isotherm of nitrogen at liquid nitrogen temperature using Quantachrome NovaWin2 version 9.0.

Fig. 2 gives the evolution of the surface area of catalysts prepared in acid pH (pH= 4.5) for the three series of catalysts.

Fig. 3 gives the evolution of the surface area of catalysts prepared in basic pH (pH= 10.5) for the three series of catalysts.

From the results, it can be observed that the specific surface area decreased with increasing molybdenum and tungsten loading.

Based on curves of Fig. 2 and Fig. 3 the surface area of serie 3 (catalysts containing Mo and W) are higher than catalysts of serie 2 (catalysts containing w) and serie 1 (catalysts containing Mo)

Fig. 2 and 3 show that the surfaces area obtained in basic pH (pH= 10.5) are higher then obtained in acid pH; this result was explained by I.M.KOLESNIKOV and coll [26, 27] by relative speeds of hydrolysis and condensation reactions during the gelation. In acid pH the speed of condensation is very low compared with hydrolysis reaction, the prepared gel was not very ramified and the resulting oxide will be with low surface area.

The surface area of catalysts is inversely proportional to the metal contained in the catalyst. This fact could be interpreted by the formation of octahedral molybdates species (or tungstates octahedral in the tungsten case) those could block the catalyst pores justifying the reduction in surface area [28, 29].

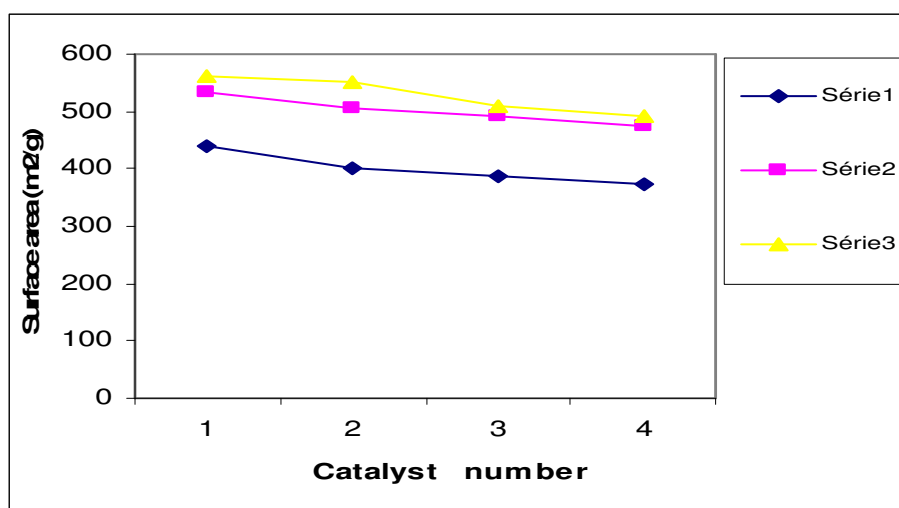


Figure 2. Surfaces area at pH = 4.5

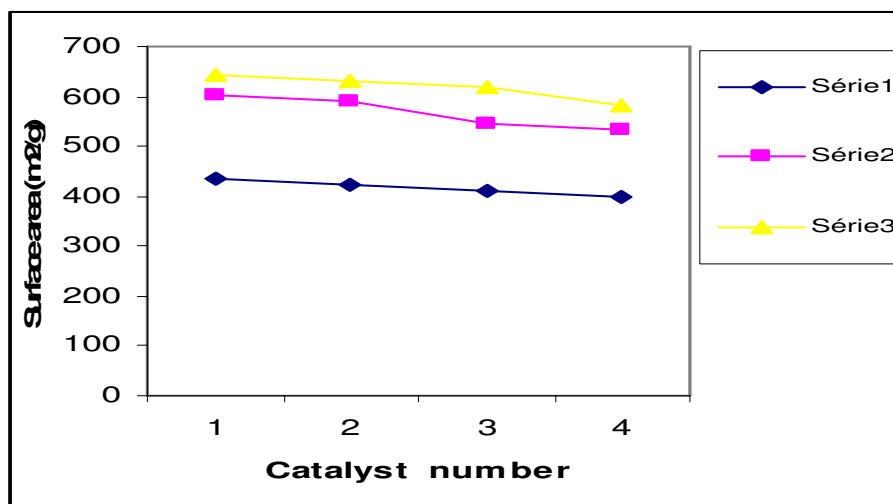


Figure 3. Surfaces area at pH = 10.5

2.2.3 Porous volume

Fig. 4 gives the evolution of the porous volume of catalysts prepared in acid pH (pH= 4.5) for the three series of catalysts.

Fig. 5 gives the evolution of the porous volume of catalysts prepared in basic pH (pH= 10.5) for the three series of catalysts.

The results show that the porous volume of catalysts vary with the pH of preparation. The catalysts with pH = 10.5 have a porous volumes higher than those obtained with catalyst at pH = 4.5, this result is related to the micellar organization of colloidal particles during the preparation of sols.

The increasing on the metal content causes a decrease in porous volume, this result is related to the formation of octahedral molybdates (tungstates), which could lead to blocking of micropores [30, 31].

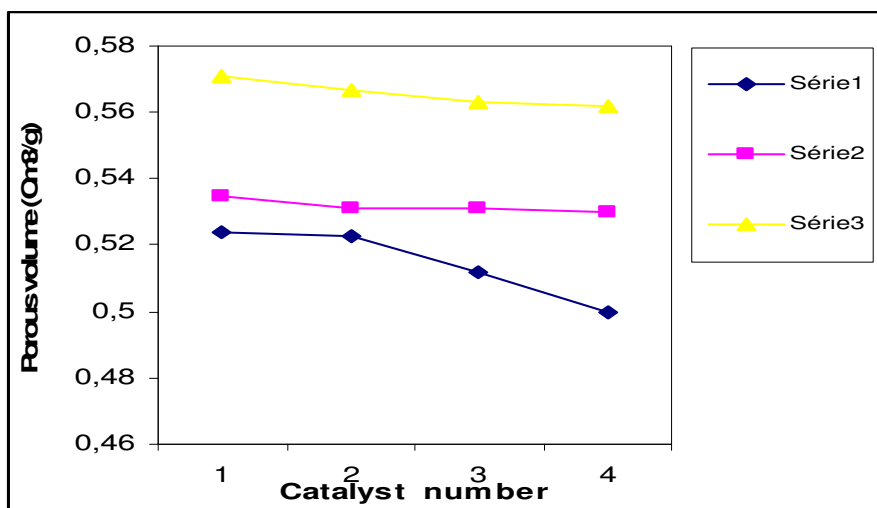


Figure 4. Porous volume at pH = 4.5

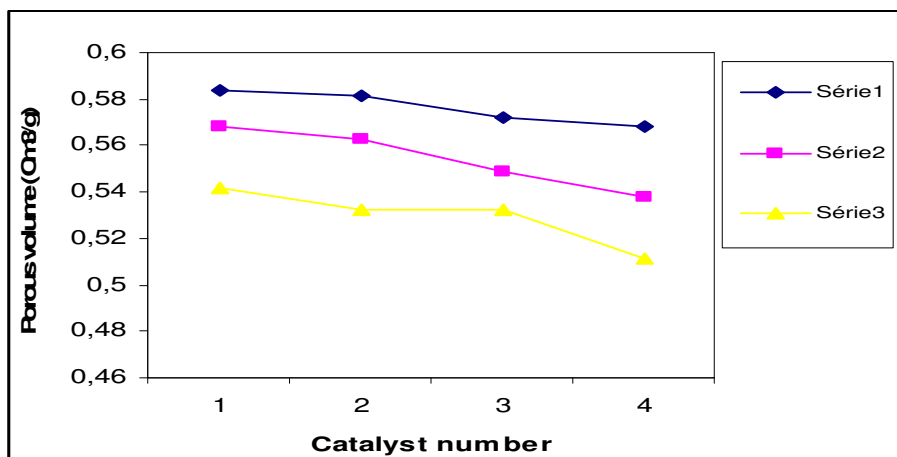


Figure 5. Porous volume at pH = 10.5

III. CATALYTIC TESTS

In order to evaluate the catalytic performances of prepared catalysts, we have carried out the catalytic tests on the isomerization of light naphtha (C_5 , C_6) obtained by the crude oil distillation unit.

3. 1. Light Naphtha Characteristics

Light naphtha (C_5 , C_6) characteristics are mentioned in table 4 and 5.

Table 4. Naphtha characteristics

Characteristics	Naphtha
Density at 15 °C	0.6496
Vapor tension (KPa)	75
Index of refraction at 20 °C	1.3740
Octane number (NOR)	64
Copper foil corrosion	negative

Table 5. Chromatographic analysis of naphtha

Family	% weight
Normal paraffins	53.390
Iso-paraffins	32.493
Naphtenes	11.298
Aromatic	2.668
Unknown	0.151

3. 2. Operating Conditions

The operating conditions of catalytic tests are mentioned in table 6.

Table 6. Operating Conditions

Parameter	Value
Temperature	250°C
Weight hourly space velocity	2 h ⁻¹
Hydrogen Pressure	8 atm
Mass of catalyst	2 g
Grains diameter	0.5 mm

The catalytic experiments were performed in a continuous flow fixed-bed quartz reactor, loaded with 2 g of the catalyst, operated under isothermal conditions, and heated by a controlled temperature electrical oven. The naphtha conversion was studied at atmospheric pressure, 250 C, weight hourly space velocity (WHSV) = 2 h⁻¹.

3. 3. Analyze of obtained products (isomerase) by chromatography

The chromatographic analyses (table 7) of obtained products are realized on chromatograph Hewlett Packard (modèle 6890 N).

Table7. Chromatographic Analyses of Obtained Products

Product obtained by Catalyst (isomerase) :	% weight				
	n-paraffins	Iso-paraffins	Naphtene	Aromatic	Unknown
(MoO ₃) _{1A}	42.413	46.139	9.312	1.964	0.172
(MoO ₃) _{2A}	40.893	48.004	9.107	1.846	0.150
(MoO ₃) _{1B}	38.408	53.324	6.505	1.632	0.151
(MoO ₃) _{2B}	36.225	54.193	7.621	1.782	0.179
(WO ₃) _{1A}	39.780	48.193	8.713	2.136	1.178
(WO ₃) _{2A}	37.391	52.723	9.054	1.607	0.161
(WO ₃) _{1B}	36.225	54.193	7.621	1.782	0.179
(WO ₃) _{2B}	33.542	57.593	6.432	1.932	0.501
(Mo-WO ₃) _{1A}	36.912	52.803	8.116	1.993	0.176
(Mo-WO ₃) _{2A}	34.391	57.093	7.661	0.734	0.121
(Mo-WO ₃) _{1B}	34.545	57.493	6.727	1.083	0.152
(Mo-WO ₃) _{2B}	27.891	63.293	6.644	1.559	0.613

3. 4 Octane Number Determination of Obtained Products (Isomerisat)

Octane number (table 8) of obtained products (isomerase) was determined by using chromatographic analyses and calculating method of Walsh [32, 33].

Table 8. Octane number of isomerase

Product obtained by Catalyst (isomerase)	Octane number (NOR)
(MoO ₃) _{1A}	70
(MoO ₃) _{2A}	71
(MoO ₃) _{1B}	73
(MoO ₃) _{2B}	73
(WO ₃) _{1A}	71
(WO ₃) _{2A}	72
(WO ₃) _{1B}	74
(WO ₃) _{2B}	80
(Mo-WO ₃) _{1A}	73
(Mo-WO ₃) _{2A}	81
(Mo-WO ₃) _{1B}	82
(Mo-WO ₃) _{2B}	83

Chromatographic analysis of obtained products shows a considerable increase of isoparaffins content compared to the naphtha, these results confirms the isomérisantes properties of prepared catalysts.

The obtained products by catalysts prepared in basic pH gave high content of isoparaffins compared to those prepared in acid pH.

The simultaneous presence of molybdenum and tungsten in the catalyst improved the isoparaffins content compared to monometallic catalysts; this result is dependant on the bimetallic character of catalysts and a synergy between two metals for the isomerisation reaction.

The isoparaffins content is higher when increased the metal quantity in the catalyst; thus increased metal function of the catalysts prompts the formation of isomers via a reduction of the diffusion path between two metallic sites. Hence, the possibility that the intermediate species would encounter acid sites and would be cracked during its migration from one metallic site to another is also less [34].

IV. CONCLUSIONS

The sol-gel method allow to prepare molybdenum and tungsten catalysts for light naphtha isomerization, the obtained results show that surface area is according to the quantity of metal contained in catalyst. The surface area obtained in basic pH is higher than those obtained in acid pH.

The products obtained (isomérate) by the light naphtha (C₅ - C₆) isomerization with the presence of prepared catalysts present isoparaffins contents and octane numbers higher compared to the naphtha.

The addition of obtained products (Isomérate) into the pool gasoline allows on the one hand, to obtain gasolines with high octane numbers, witch respect the international and environmental requirements and on the other hand, the possibility of developing the light fractions (C₅ - C₆) resulting from the crude oil distillation unit and condensed split distillation.

REFERENCES

- [1] M.H. Jordao, V. Simoes, D. Cardoso, Appl. Catal. 319 (2007) 1.
- [2] S.D. Rossi, G. Ferraris, M. Valigi, D. Gazzoli, Appl. Catal. 231 (2002)173.
- [3] S. Zhang, Y. Zhang, J.W. Tierney, I. Wender, Appl. Catal. A 193 (2000) 155.
- [4] K. Watanabe, T. Kawakami, K. Baba, N. Oshio, T. Kimura, Catal. Surv. Asia 9 (2005)17.
- [5] P.J. Kucher, J.C. Bricker, M.E. Reno, R.S. Haizmann, Fuel Process. Technol. 35 (1993)183.
- [6] S. Kuba, M. Che, R.K. Grasselli, H. Knozinger, J. Phys. Chem. B 107 (2003) 3459.
- [7] S.G. Ryu, B.C. Gates, Ind. Eng. Chem. Res. 37 (1998) 1786.
- [8] N. Essayem, Y. Ben Taa^rit, C. Feche, P.Y. Gayraud, G. Sapaly, C. Naccache, J. Catal. 219 (2003) 97.
- [9] P.Lepince, « Procédés de transformation », Editions Technip, 1998, Paris
- [10] V.M. Ben.tetz, C.R. Vera, C.L. Pieck, F.G. Lacamoire, J.C. Yori, J.M. Grau, J.M. Parera, Catal. Today 107– 108 (2005) 651.
- [11] M.Belloum, Thèse de doctorat. Université Piere et Marie Curie. (1990)
- [12] Belatel H, Al-Kandari H, Al-Khorafi F, Katrib A., Garin F., Appl. Catal . A: General 275(2004)141.
- [13] Y.Holl, F.Garin.Maire, P.A.Engelhard and J.Grosmangin, Appl.Catal.46(1989)57
- [14] H. Hino, K. Arata, J.Chem, Soc, Chem. Commun. (1987)1259
- [15] E. Iglesia, J. E. Baumgartner, F. H. Ribeiro and M. Boudart J. Catal.,131, iss.2 (1991) 523
- [16] F. H. Ribeiro, R. A. D. Betta, M. Boudart, J. Catal., 130 (1991) 86.
- [17] Phalipon . J. une nouvelle voie de synthèse d'oxydes : Applications aux verres. L'industrie céramique, 1987, n 81, n 2, p11 -18
- [18] Pierre A. C, Introduction aux procédés sol-gel, Paris : SEPTIMA, 1992, 199
- [19] BRINKER. C. J, SHERER.G.w, « Sol gel science: The physics and chemistry of
- [20] J. Livage, Matériaux et Techniques N 6-7 (1994)
- [21] M. Campanati, G. Fornasari, A. Vaccari, Catal. Today 77 (2003) 299.
- [22] F. Schmidt, Appl. Catal. 221 (2001) 15.

- [23] M. Campanati, G. Fornasari, A. Vaccari, Catal. Today 77 (2003) 299.
- [24] F. Schmidt, Appl. Catal. 221 (2001) 15.
- [25] I.M.Kolesnikov, A.V.Yuablousky, S.I.Kolesnikov, A. Busenna et M.Y.Kiljanov, « Preparation of Metallo-Silicate Solid Satalysts by Sol-gel method with Regulation of Activity and Selectivity” 6ème International Symposium, Vol 3, session 2 sept. 15-1 (1994)
- [26] C.L.Glaves, C.J.Brinker; D.M.Smith; and P.J.Davis. Chemistry of Materials 1989, 1,1
- [27] T.N.Vu; J. Van Gestel, J.P.Gilson; C.Collet, J.P.Dath; J.C.Duchet. “Platinum tungstated Zirconia isomérisation catalysts; Part 1: Characterization of Acid and Metal propreties”. Journal of Catalyst 231 (2005)453-467
- [28] R. Lopez-Cordero; F.J.Gil L lambias; A. Lopez Agudo, Appl. Catal. A74 (1991).
- [29] R. Lopez-Cordero; F.J.Gil L lambias; A. Lopez Agudo, Appl. Catal. A74 (1991) 125.
- [30] M. Signoretto, M. Scarpa, F. Pinna, G. Strkul, P. Canton, A. Benedetti, J. Non-Cryst. Solids 225 (1998) 178.
- [31] T.N.Vu; J. Van Gestel, J.P.Gilson; C.Collet, J.P.Dath; J.C.Duchet. “Platinum tungstated Zirconia isomérisation catalysts; Part 1: Characterization of acid and metal propreties”. Journal of catalyst 231 (2005)453-467
- [32] Guemini M, Rezgui. Y, “Isomérisation du n-hexane sur des Catalyseurs Ni-WO_x / Al₂O₃-SiO₂ ». The Canadian Journal of Chemical Engineering, Volume 82, February 2004.
- [33] N.Petroff « Automatic Determination of Reformate Octane Number and Composition by Gas Chromatography» Oil and Gas Science and Technology- Rev IFP, Vol 43 (1988), No 2
- [34] R.P.Walsh, J.V.Mortimer «New Way to Test product quality», Hydrocarbon Processing, 1971, 153

Authors

Farid Aoudjit has received Master's Degree in applied chemistry, specialization adsorbents and catalysts, Faculty of hydrocarbons and chemistry, university of Bomerdes. Doctorat student (2009 until now). Fields of interests: adsorption, catalysts and zeolite.



MEMBERS OF IJAET FRATERNITY

Editorial Board Members from Academia

Dr. P. Singh,

Ethiopia.

Dr. A. K. Gupta,

India.

Dr. R. Saxena,

India.

Dr. Natarajan Meghanathan,

Jackson State University, Jackson.

Dr. Rahul Vaish,

School of Engineering, IIT Mandi, India.

Dr. Syed M. Askari,

University of Texas, Dallas.

Prof. (Dr.) Mohd. Husain,

A.I.E.T, Lucknow, India.

Dr. Vikas Tukaram Humbe,

S.R.T.M University, Latur, India.

Dr. Mallikarjun Hangarge,

Bidar, Karnataka, India.

Dr. B. H. Shekar,

Mangalore University, Karnataka, India.

Dr. A. Louise Perkins,

University of Southern Mississippi, MS.

Dr. Tang Aihong,

Wuhan University of Technology, P.R.China.

Dr. Rafiqul Zaman Khan,

Aligarh Muslim University, Aligarh, India.

Dr. Abhay Bansal,

Amity University, Noida, India.

Dr. Sudhanshu Joshi,

School of Management, Doon University, Dehradun, India.

Dr. Su-Seng Pang,

Louisiana State University, Baton Rouge, LA, U.S.A.

Dr. Avanish Bhadauria,

CEERI, Pilani, India.

Dr. Dharma P. Agrawal

University of Cincinnati, Cincinnati.

Dr. Rajeev Singh

University of Delhi, New Delhi, India.

Dr. Smriti Agrawal

JB Institute of Engineering and Technology, Hyderabad, India

Prof. (Dr.) Anand K. Tripathi

College of Science and Engg., Jhansi, UP, India.

Prof. N. Paramesh

University of New South Wales, Sydney, Australia.

Dr. Suresh Kumar

Manav Rachna International University, Faridabad, India.

Dr. Akram Gasmelseed

Universiti Teknologi Malaysia (UTM), Johor, Malaysia.

Dr. Umesh Kumar Singh

Vikram University, Ujjain, India.

Dr. A. Arul Lawrence Selvakumar

Adhiparasakthi Engineering College, Melmaravathur, TN, India.

Dr. Sukumar Senthilkumar

Universiti Sains Malaysia, Pulau Pinang, Malaysia.

Dr. Saurabh Pal

VBS Purvanchal University, Jaunpur, India.

Dr. Jesus Vigo Aguiar

University Salamanca, Spain.

Dr. Muhammad Sarfraz

Kuwait University, Safat, Kuwait.

Dr. Xianbo Qui

Xiamen University, P.R. China.

Dr. C. Y. Fong

University of California, Davis.

Prof. Stefanos Gritzalis

University of the Aegean, Karlovassi, Samos, Greece.

Dr. Hong Hu

Hampton University, Hampton, VA, USA.

Dr. Donald H. Kraft

Louisiana State University, Baton Rouge, LA.

Dr. Veeresh G. Kasabegoudar

COEA, Maharashtra, India.

Dr. Nouby M. Ghazaly

Anna University, Chennai, India.

Dr. Paresh V. Virparia

Sardar Patel University, V V Nagar, India.

Dr. Vuda Srinivasarao

St. Mary's College of Engg. & Tech., Hyderabad, India.

Dr. Pouya Derakhshan-Barjoei

Islamic Azad University, Naein Branch, Iran.

Dr. Sanjay B. Warkad

Priyadarshini College of Engg., Nagpur, Maharashtra, India.

Dr. Pratyosh Shukla

Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India.

Dr. Mohamed Hassan Abdel-Wahab El-Newehy

King Saud University, Riyadh, Kingdom of Saudi Arabia.

Dr. K. Ramani

K.S.Rangasamy College of Tech., Tiruchengode, T.N., India.

Dr. J. M. Mallikarjuna

Indian Institute of Technology Madras, Chennai, India.

Dr. Chandrasekhar

Dr. Paul Raj Engg. College, Bhadrachalam, Andhra Pradesh, India.

Dr. V. Balamurugan

Einstein College of Engineering, Tirunelveli, Tamil Nadu, India.

Dr. Anitha Chennamaneni

Texas A&M University, Central Texas, U.S.

Dr. Sudhir Paraskar

S.S.G.M.C.E. Shegaon, Buldhana, M.S., India.

Dr. Hari Mohan Pandey

Middle East College of Information Technology, Muscat, Oman.

Dr. Youssef Said

Tunisie Telecom / Sys'Com Lab, ENIT, Tunisia.

Dr. Mohd Nazri Ismail

University of Kuala Lumpur (UniKL), Malaysia.

Dr. Gabriel Chavira Juárez

Autonomous University of Tamaulipas, Tamaulipas, Mexico.

Dr. Saurabh Mukherjee

Banasthali University, Banasthali, Rajasthan, India.

Prof. Smita Chaudhry

Kurukshetra University, Kurukshetra, Haryana, India.

Dr. Raj Kumar Arya

Jaypee University of Engg. & Tech., Guna, M. P., India.

Dr. Prashant M. Dolia

Bhavnagar University, Bhavnagar, Gujarat, India.

Dr. Sarita Azad

IIT Mandi, India.

Dr. Dewan Muhammad Nuruzzaman

Dhaka University of Engg. and Tech., Gazipur, Bangladesh.

Dr. Hadj. Hamma Tadjine

IAV GmbH, Germany.

Dr. D. Sharmila

Bannari Amman Institute of Technology, Sathyamangalam, India

Editorial Board Members from Industry/Research Labs.

Tushar Pandey,

STEricsson Pvt Ltd, India.

Ashish Mohan,

R&D Lab, DRDO, India.

Amit Sinha,

Honeywell, India.

Tushar Johri,

Infosys Technologies Ltd, India.

Dr. Om Prakash Singh ,

Manager, R&D, TVS Motor Company, India.

Dr. B.K. Sharma

Northern India Textile Reserch Assoc., Ghaziabad, U.P., India.

Advisory Board Members from Academia & Industry/Research Labs.

Prof. Andres Iglesias,

University of Cantabria, Santander, Spain.

Dr. Arun Sharma,

K.I.E.T, Ghaziabad, India.

Prof. Ching-Hsien (Robert) Hsu,

Chung Hua University, Taiwan, R.o.C.

Dr. Himanshu Aggarwal,

Punjabi University, Patiala, India.

Prof. Munesh Chandra Trivedi,

CSEDIT School of Engg., Gr. Noida, India.

Dr. P. Balasubramanie,

K.E.C., Perundurai, Tamilnadu, India.

Dr. Seema Verma,

Banasthali University, Rajasthan, India.

Dr. V. Sundarapandian,

Dr. RR & Dr. SR Technical University, Chennai, India.

Mayank Malik,

Keane Inc., US.

Prof. Fikret S. Gorgen,

Bogazici University Istanbul, Turkey.

Dr. Jiman Hong

Soongsil University, Seoul, Korea.

Prof. Sanjay Misra,

Federal University of Technology, Minna, Nigeria.

Prof. Xing Zuo Cheng,

National University of Defence Technology, P.R.China.

Dr. Ashutosh Kumar Singh

Indian Institute of Information Technology Allahabad, India.

Dr. S. H. Femmam

University of Haute-Alsace, France.

Dr. Sumit Gandhi

Jaypee University of Engg. & Tech., Guna, M. P., India.

Dr. Hradyesh Kumar Mishra

JUET, Guna, M.P., India.

Dr. Vijay Harishchandra Mankar

Govt. Polytechnic, Nagpur, India.

Prof. Surendra Rahamatkar

Nagpur Institute of Technology, Nagpur, India.

Dr. B. Narasimhan

Sankara College of Science And Commerce, Coimbatore, India.

Dr. Abbas Karimi

Islamic Azad University, Arak Branch, Arak, Iran.

Dr. M. Munir Ahamed Rabbani

Qassim University, Saudi Arabia.

Dr. Prasanta K Sinha

Durgapur Inst. of Adva. Tech. & Manag., Durgapur, W. B., India.

Dr. Tole H. Sutikno

Ahmad Dahlan University(UAD), Yogyakarta, Indonesia.

Dr. Anna Gina Perri

Politecnico di Bari, BARI - Italy.

Prof. Surendra Rahamatkar

RTM Nagpur University, India.

Dr. Sagar E. Shirsath

Vivekanand College, Aurangabad, MS, India.

Dr. Manoj K. Shukla

Harcourt Butler Technological Institute, Kanpur, India.

Dr. Fazal Noorbasha

KL University, Guntur, A.P., India.

Dr. Manjunath T.C.

HKBK College of Engg., Bangalore, Karnataka, India.

Dr. M. V. Raghavendra

Swathi Institute of Technology & Sciences, Ranga Reddy , A.P. , India.

Research Volunteers from Academia

Mr. Ashish Seth,

Ideal Institute of Technology, Ghaziabad, India.

Mr. Brajesh Kumar Singh,

RBS College, Agra, India.

Prof. Anilkumar Suthar,

Kadi Sarva Viswavidhayalay, Gujarat, India.

Mr. Nikhil Raj,

National Institute of Technology, Kurukshetra, Haryana, India.

Mr. Shahnawaz Husain,

Graphic Era University, Dehradun, India.

Mr. Maniya Kalpesh Dudabhai

C.K.Pithawalla College of Engg.& Tech., Surat, India.

Dr. M. Shahid Zeb

Universiti Teknologi Malaysia(UTM), Malaysia.

Mr. Brijesh Kumar

Research Scholar, Indian Institute of Technology, Roorkee, India.

Mr. Nitish Gupta

Guru Gobind Singh Indraprastha University, India.

Mr. Bindeshwar Singh

Kamla Nehru Institute of Technology, Sultanpur, U. P., India.

Mr. Vikrant Bhateja

SRMGPC, Lucknow, India.

Mr. Ramchandra S. Mangrulkar

Bapurao Deshmukh College of Engineering, Sevagram, Wardha, India.

Mr. Nalin Galhaut

Vira College of Engineering, Bijnor, India.

Mr. Rahul Dev Gupta

M. M. University, Mullana, Ambala, India.

Mr. Navdeep Singh Arora

Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India.

Mr. Gagandeep Singh

Global Institute of Management and Emerging Tech., Amritsar, Punjab, India.

Ms. G. Loshma

Sri Vasavi Engg. College, Pedatadepalli, West Godavari, Andhra Pradesh, India.

Mr. Mohd Helmy Abd Wahab

Universiti Tun Hussein ONN Malaysia, Malaysia.

Mr. Md. Rajibul Islam

University Technology Malaysia, Johor, Malaysia.

Mr. Dinesh Sathyamoorthy

Science & Technology Research Institute for Defence (STRIDE), Malaysia.

Ms. B. Neelima

NMAM Institute of Technology, Nitte, Karnataka, India.

Mr. Mamilla Ravi Sankar

IIT Kanpur, Kanpur, U.P., India.

Dr. Sunusi Sani Adamu

Bayero University, Kano, Nigeria.

Dr. Ahmed Abu-Siada

Curtin University, Australia.

Ms. Shumos Taha Hammad

Al-Anbar University, Iraq.

Mr. Ankit R Patel

L C Institute of Technology, Mahesana, India.

Mr. Athar Ravish Khan Muzaffar Khan

Jawaharlal Darda Institute of Engineering & Technology Yavatmal, M.S., India.

Prof. Anand Nayyar

KCL Institute of Management and Technology, Jalandhar, Punjab, India.

Mr. Arshed Oudah

UTM University, Malaysia.

Mr. Piyush Mohan

Swami Vivekanand Subharti University, Meerut, U.P., India.

Mr. Mogaraju Jagadish Kumar

Rajampeta, India.

Mr. Deepak Sharma

Swami Vivekanand Subharti University, Meerut, U.P., India.

Mr. B. T. P. Madhav

K L University, Vaddeswaram, Guntur DT, AP, India.

Mr. Nirankar Sharma

Subharti Institute of Technology & Engineering, Meerut, U.P., India.





Volume-2, Issue-1

URL : <http://www.ijaet.org>